

71742

STIC-Biotech/ChemLib

From: Mehta, Ashwin
Sent: Wednesday, July 24, 2002 2:49 PM
T : STIC-Biotech/ChemLib
Subject: seq search

STIC,

Please search the commercial and interference databases for the amino acid seq of SEQ ID NO: 1 and the nucleotide sequence of SEq ID NO: 2 from 09/971,020.

My mail room is 9E12, office 9E07, art unit 1638.

Thank you,
Ashwin

Ashwin Mehta
United States Patent and Trademark Office
Biotechnology Patent Examiner
703-306-4540

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 8/25/02
Date Completed: 8/28/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: OP
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 27, 2002, 03:27:24 ; Search time 242.08 Seconds
(without alignments)
9205.861 Million cell updates/sec

Title: US-09-971-020-2

Perfect score: 1298
Sequence: 1 agcagtcgcaatcgatgtg.....atnaaaaaaaaaaaaaa 1298

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

N.Geneseq.032802.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138.8	10.7	1427	22	AAC82742
2	138.8	10.7	1427	22	AAC82743
3	134.8	10.4	1363	22	AAF80136
4	129.8	10.0	1407	21	AAC35921
5	126.8	9.8	1239	21	AAC45306
6	90.4	7.0	1170	24	ABA05043
7	90.4	7.0	1345	21	AAC30544
8	90.4	7.0	1476	24	ABA05044
9	50.4	3.9	1206	21	AAC46773

10	48.8	3.8	1211	21	AAC33567	Arabidopsis thaliana
11	43.8	3.4	128	22	AA07727	Cervical cancer pr
12	42.6	3.3	1259	21	AAC48777	Arabidopsis thaliana
13	42.6	3.3	1261	21	AAC36551	Arabidopsis thaliana
14	41.4	3.2	3637	20	AAK60802	Human secreted pro
15	41.4	3.2	3637	22	AAK59208	Human secreted pro
16	41.4	3.2	3637	24	ABA90877	Human cDNA encodin
17	41.2	3.2	5928	24	ABL34087	Human polynucleoti
18	41	3.2	5928	24	ABL34087	Human immune syste
19	41	3.2	873	24	AAAD22204	Renilla reniformis
20	41	3.2	1299	18	AAAT73692	DNA encoding aspar
21	41	3.2	1299	19	AAV44040	S. aureus asparagi
22	41	3.2	14041	22	AAH48024	Internal control B
23	40.8	3.1	5294	24	ABL33071	Human immune syste
24	40.6	3.1	612	22	AAH71471	Cervical cancer pr
25	40.6	3.1	612	22	AAH71471	Human cervical can
26	40.6	3.1	623	21	AAC33628	Arabidopsis thaliana
27	40.6	3.1	711	22	AAAL5484	Human breast cance
28	40.6	3.1	5884	24	ABL34164	Human immune syste
29	40.4	3.1	15247	24	ABL34083	Human immune syste
30	40	3.1	425	22	AAK60450	Human cancer agent
31	39.8	3.1	3839	21	AAAD26336	Human secreted pro
32	39.8	3.1	6365	24	ABL32124	Human immune syste
33	39.8	3.1	9965	24	ABL33527	Human immune syste
34	39.8	3.1	56152	22	AAK46793	Tumour suppressor
35	39.6	3.1	6185	24	ABL32413	Human immune syste
36	39.6	3.1	19734	24	ABL33932	Human immune syste
37	39.4	3.0	760	22	ABA83286	Human secreted pro
38	39.4	3.0	1830	20	AAK00612	Human secreted pro
39	39.4	3.0	1830	21	AAK98100	Human secreted pro
40	39.4	3.0	1830	22	AAAD1713	Human secreted pro
41	39.4	3.0	1842	21	AAK98096	Human secreted pro
42	39.4	3.0	1842	22	AAAD1709	Human secreted pro
43	39.4	3.0	2218	20	AAV84503	Human secreted pro
44	39.4	3.0	2218	22	ABA83415	Human secreted pro
45	39.4	3.0	6311	24	ABL33963	Human immune syste

ALIGNMENTS

RESULT 1	
AC82742	standard; DNA; 1427 BP.
AC82742	
AC82742	
16-MAR-2001 (first entry)	
C. sinensis N-methyl transferase DNA.	
Caffeine biosynthesis; N-methyl transferase; tea; N-met; plant; ds.	
Camellia sinensis.	
EP1055727-A2.	
29-NOV-2000.	
26-MAY-2000; 2000EP-0304522.	
26-MAY-1999; 99JP-0146358.	
(MITA) MITSUI CHEM INC.	
Mizuno M, Ashihara H, Mizuno K, Fujimura T;	
WPI: 2001-063619/08.	
P-PSDB; AAB45799.	
DNA encoding N-methyl transferase with enzyme activities of 7-methyl	
xanthine N3 methyl transferase, theobromine N1 methyl transferase and	
paraxanthine N3 methyl transferase, useful for caffeine synthesis in	

[illegible]

	XX	Key	Location/Qualifiers
	XX	FT CDS	19..113
	XX	FT	/tag- a
	XX	FT	/product= "benzoic acid carboxyl methyltransferase"
	XX	PX	MO200118183-A1.
	XX	PX	15-MAR-2001.
	XX	PF	01-SEP-2000; 2000OMO-US24362.
	XX	PX	03-SEP-1999; 99OUS-0152393.
	XX	PA	(PURD) PURDUE RES FOUND.
	XX	PI	Doudareva N, Murfitt LM, Mann CJ;
	XX	PI	MP1: 2001-244568/25.
	XX	P-PSDB:	AAB67769.
	XX	PT	Newel purified benzoic acid carboxyl methyltransferase protein useful
	XX	PT	for production of floral scent compounds and biosynthesis -
	XX	PS	Claim 9; Page 85-88; 94pp: English.
	XX	CC	The present sequence encodes a Snapdragon benzoic acid carboxyl
	XX	CC	methyltransferase (BAMT) protein. The protein is a
	XX	CC	S-adenosyl-L-methionine:benzoic acid carboxyl methyltransferase, and
	XX	CC	the enzyme is used in the formation of floral scent compounds such as
	XX	CC	methyl benzoate. The enzyme catalyzes the transfer of a methyl group
	XX	CC	donor, such as S-adenosyl methionine (SAM) to benzoic acid to form
	XX	CC	methyl benzoate. BAMT proteins are useful for the production of floral
	XX	CC	scent compounds and biosynthesis. BAMT polynucleotides are useful for
	XX	CC	producing transgenic plants.
	XX	Sequence	1363 BP; 417 A; 239 C; 300 G; 407 T; 0 other;
	XX	Query Match	10.4%; Score 134.8; DB 22; Length 1363;
	XX	Best Local Similarity	49.6%; Pred. NO. 3.4e-26;
	XX	Matches 546; Conservative	0; Mismatches 482; Indels 72; Gaps
Oy	59	aatgaaggttaagcgatacagaacttcggcaagaatgcatactcgtgcttgc	118
Oy	52	atcgagagagatggttgaaacactgctacgccacaattcggcttcaaaaagtatg	111
Oy	119	aaggatgaaccttcccttgaacaatcatcagaaatgtgtcgggccaattggccaac	178
Oy	112	tcaaatcatgtcatgatttttagagaagaacccttaagaatatcatatcgtgatgtgc	171
Oy	179	atcaaagaatgcataaagttgvcggaattggnatgccttcgtgaccaaacacacttta	238
Oy	172	ttcccanaatgcctcaaatgatafgatagtgattgtttcatcacggatcagggcctttg	231
Oy	239	acagatgcggagcattgttcgaagatgtgacaagaattgcgcagagaagaagaatgaat	298
Oy	232	gtcatgctcgcgactataataacatttggagttt---gtacaacgagaagaattpat	288
Oy	299	gaagctccacaccttcagatttttcgataatcattttcaaaatgatttcaattcggtt	358
Oy	289	gaattactgaatttgttggttttttcgaagcatcttcagacaagcttcaacaactc	348
Oy	359	ttaaagttgttgcgaagctcttacacgcaactcagagaagaataatgagcaggaatga	418
Oy	349	tccaattgttatc-----aatatgagaatiga	375
Oy	419	tatgagctaataagcgaatgctgctcttctcaaggcagagctcttccccgaatgattcc	478
Oy	376	aactgtcttataatgattcttgcgcgaatcttcttcaaggagactatgtgcaaaaaagc	435
Oy	479	atgaatttttgcactctgttacaagtgatcaattgttatctcaagttcccaaggatttg	538

[illegible]

Db	701	aagtcatacaaggccactgaatcatcaattccaaaagaagcttcacacatggttttaaggt	760	PR	23-APR-1999	99US-0130851	
Oy	661	tcatttcgaagaagttgttttcacgtgcgcgaagtcgaccttacttactgtttaaagtaga	720	PR	28-APR-1999	99US-0131449	
Db	761	acgtctcgaagaattgtctctcaatgacgcatggtctccactctcatcggaagaacac	820	PR	30-APR-1999	99US-0132048	
Oy	721	tgaattcgagcaaccgaatccctcagact-----tacttgatgagaaataa-----	768	PR	04-MAY-1999	99US-0132464	
Db	821	tcctaacgataccatgtatagaagatgtgttctactcttcttgaaatgtcatcaacctct	880	PR	05-MAY-1999	99US-0132465	
Oy	769	---gacttgatgttgtagagactcttgagagaagaataatgtatagttcattcttc	825	PR	06-MAY-1999	99US-0132467	
Db	881	tcgtgacctatctcttgagagctctctgtgagtgatcaaaagctgacgcatcaatgccc	940	PR	07-MAY-1999	99US-0132468	
Oy	826	attcttcaccttcagcagagaagaaagtaagtgatcatgttgagagaaggtcttcgga	885	PR	11-MAY-1999	99US-0134256	
Db	941	gtttatgatacccgaaagcagtaacaagaaactaaagaagtgatacaaaaagaggtccttga	1000	PR	14-MAY-1999	99US-0134257	
Oy	886	aatttatactctgagacattttaaagccattatgactgctcctctctatgtatgata	945	PR	14-MAY-1999	99US-0134258	
Db	1001	aaacaatgaattgag---tcacatgattgttgaactgtgtccctctacgaagaagatba	1057	PR	20-MAY-1999	99US-0135124	
Oy	946	ttacccgaatgaatcccatgataaacaataaagacagagatgtgtgatacttaattagatc	1005	PR	21-MAY-1999	99US-0135125	
Db	1058	cttgtaagcagaagcagcaatgaa-----gctaatgacataagac	1096	PR	24-MAY-1999	99US-0135629	
Oy	1006	agtttcaacaaccctccctgcagacatcttcttgagaagcatatcatgcctcatcttca	1065	PR	25-MAY-1999	99US-0135633	
Db	1097	tgttatgtacacaatgctctctatgttcattcttgagagaatcatgatacctgttcca	1156	PR	28-MAY-1999	99US-0136392	
Oy	1066	cagagcttcgagaagatgcagcaaaagttctccacaatgggcaaggtcgtcatataatct	1125	PR	07-JUN-1999	99US-0137524	
Db	1157	taagtgtacatacatctgtgcttcaacaatgacgaacatgcaggaacaaagcgttcagct	1216	PR	08-JUN-1999	99US-0138094	
Oy	1126	tatcatctctctgcgcaaaagcaggaagtcagacgtct	1166	PR	10-JUN-1999	99US-0138847	
Db	1217	tgctgttcttcgtgactaagaagtaatcaactctgcacgt	1217	PR	16-JUN-1999	99US-0139452	
RESULT 5					PR	17-JUN-1999	99US-0139452
AAC45306					PR	18-JUN-1999	99US-0139454
AAC45306 standard; DNA; 1239 BP.					PR	18-JUN-1999	99US-0139455
AAC45306;					PR	18-JUN-1999	99US-0139456
18-OCT-2000 (first entry)					PR	21-JUN-1999	99US-0139817
Arabidopsis thaliana DNA fragment SEQ ID NO: 46034.					PR	22-JUN-1999	99US-0139899
Hybridisation assay; genetic mapping; gene expression control;					PR	23-JUN-1999	99US-0140353
protein identification; signal transduction pathway;					PR	23-JUN-1999	99US-0140354
metabolic pathway; promoter; termination sequence; ss;					PR	24-JUN-1999	99US-0140655
Arabidopsis thaliana.					PR	28-JUN-1999	99US-0140823
EP1033405-A2.					PR	29-JUN-1999	99US-0140991
06-SEP-2000.					PR	30-JUN-1999	99US-0141287
25-FEB-2000; 2000EP-0301439.					PR	01-JUL-1999	99US-0141842
25-FEB-1999; 99US-0121825.					PR	01-JUL-1999	99US-0142154
09-MAR-1999; 99US-0123180.					PR	02-JUL-1999	99US-0142055
23-MAR-1999; 99US-0123548.					PR	06-JUL-1999	99US-0142390
25-MAR-1999; 99US-0125788.					PR	08-JUL-1999	99US-0142803
29-MAR-1999; 99US-0126264.					PR	09-JUL-1999	99US-0142920
01-APR-1999; 99US-0127462.					PR	12-JUL-1999	99US-0142977
06-APR-1999; 99US-0128234.					PR	13-JUL-1999	99US-0143542
16-APR-1999; 99US-0128714.					PR	14-JUL-1999	99US-0143624
19-APR-1999; 99US-0129845.					PR	15-JUL-1999	99US-0144005
21-APR-1999; 99US-0130							

Db 821 tctaacgacatcattgataagattgtgtcacttttgacattgctacaactctct 880
 QY 770 ----gacttgattgttgaggacttctgagagaaanaattgattgattcaattcc 825
 Db 881 ccgtgacctgacttcttgaggctcttgagatgagtcacaaagcttgacgcttcaaatgcc 940
 QY 826 attcttacccttcaagagaagaagatgagatgatttgagaggaaggtctctgca 885
 Db 941 gtttaagatccgaacgacacgaactcaagaagtgtacacaaagaggtctcttga 1000
 QY 886 aatttatatctgagacttttaagccattatgctgctcctctcattgattgagta 945
 Db 1001 aatcaatgatttgag-----tcaatgatttgattcttctgctacacgaagaagatga 1057
 QY 946 ttaccgagtagaagctccatgacacaaatgaagcagagatgtgacatcattatgaagc 1005
 Db 1058 cttggaagcagagcagcagatga-----gctaaagcagataagc 1096
 QY 1006 agttaagaaacccatccctcgaagatcatttggagaagctattatgctgacttattcca 1065
 Db 1097 tgttagtgaacaaatgctcattgtctcatttggagaagaattatcgatccttctgca 1156
 QY 1066 caggcttgcaagcagacgaagaggtctccacatggcgaaggtgctgataataatct 1125
 Db 1157 taagatagcatcactgactgactcaacatgccaactgacagaaacagacgtgctagct 1216
 QY 1126 tatcattctctcgcaaaaag 1147
 Db 1217 tgcgttctcttactgaag 1238
 RESULT 6
 ID ABA05043 standard: cDNM: 1170 BP.
 AC ABA05043;
 XX
 DT 22-FEB-2002 (first entry)
 DE A thaliana jasmonic acid carboxyl methyltransferase coding sequence #1.
 XX
 KM Jasmonic acid carboxyl methyltransferase; virus resistance;
 KM Insect resistance; cold resistance; drought resistance; antifungal;
 KM S-adenosyl-L-methionine: jasmonic acid carboxyl methyltransferase;
 KM bacterial phytopathogen resistance; salt resistance; JMT;
 KM antibacterial; antiviral; transgenic plant; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1170
 FT /tag- a
 FT /product- "jasmonic acid carboxyl methyltransferase"
 PN MO200196549-A1.
 XX
 PD 20-DEC-2001.
 XX
 PF 05-JUN-2001; 2001MO-KR00953.
 XX
 PR 13-JUN-2000; 2000KR-0032365.
 XX
 PA (SCIG-) SCIGEN HARVEST CO LTD.
 PA (CHOI/) CHOI Y.
 PI Choi Y, Cheong J, Lee J, Song J, Song S, Seo H, Koo Y;
 XX WPI, 2002-090208/12.
 DR P-Psdb; AAM47309.
 XX
 PT Jasmonic acid carboxyl methyltransferase and gene encoding it useful
 PT for enhancing a resistance of plant against damages caused by
 PT phytopathogens and harmful insects, and stresses

XX Claim 3; Page 48-49; 56pp; English.
 PS
 CC The present invention relates to the Arabidopsis thaliana jasmonic acid
 CC carboxyl methyltransferase (JMT, also known as S-adenosyl-L-methionine:
 CC jasmonic acid carboxyl methyltransferase). The sequences can be used to
 CC produce transgenic plants with increased resistance to cold, salt,
 CC drought, viruses, harmful insects and bacterial phytopathogens. The
 CC present sequence is one version of the cDNA of the invention.
 XX
 SO Sequence 1170 BP; 317 A; 244 C; 299 G; 310 T; 0 other;
 Query Match 7.08; Score 90.4; DB 24; Length 1170;
 Best Local Similarity 49.48; Pred No. 2.9e-14;
 Matches 380; Conservative 0; Mismatches 351; Indels 39; Gaps 4;
 QY 415 agatgctgcttaataaagcgaatgctgctcttctcaagcagacattcccgagga 474
 Db 393 agaatcgcttgctgctgagcgcgcgcgaaggtgcttcaagcagcttgctccgcg 452
 QY 475 gtccatgcatattttgacactctgttacagtgctcattgtgtatctcaggtccagcgg 534
 Db 433 gacgttcacttgctgcatcttctcttgatattgattgtgtgctcaggttccatgctg 512
 QY 535 ttgtgtattgaaattggggat-----tggtcaacaagaaggagattta 579
 Db 513 tgagcgaggaaggaagacagacataaagcgtgattagaacacatgggaaataata 572
 QY 580 ctctccaaagagatgctgcgcgcgcgcgcgaagagcatatttgatcaatttaagaaga 639
 Db 573 catatcaagaagacagctccttaagagtgacatcaaaagcttcttcaattcaacacga 632
 QY 640 tttaacacatttcaagatcatttgaagaaggttttcaagtcgagcgaatctct 699
 Db 633 ttcttgcttcttcttgagtcagatctgagaggtttcccggaagcagcagctgttt 692
 QY 700 taactgacttgt--aaagtatgaaattcgacacga-----atccctaga 747
 Db 693 atcgctccttgagagaagatcactgagatcccaaacgaagaagattgctataatggga 752
 QY 748 cttaactgacatgacataaagacattgattgtgaggacttctggaagaaagaaatt 807
 Db 753 actcctgctcaagctccttaagtctacagcgaagaagaggtatcatcgagaagaagaat 812
 QY 808 ggaatgctcaatcattcattcttaccacttcaagaagaagaagtaagtgatagttga 867
 Db 813 cgatgcttcaagctccttactatgtgtgagctccgaagagtgaaatgtgataga 872
 QY 866 ggaagaaaggttcttggaatttatatctggaagacttgaagccattatgagctgc 927
 Db 873 gaaagaaggtcattcttcgacgataagcttgagataagctcgatttgagaaagctgc 932
 QY 928 ctctctatgatagtatcaccagta-----agatccatgaacaataaa 975
 Db 933 gatatcagtgagagaagttatgaccttgaataaagctccaaacccgaagccctagctag 992
 QY 976 agcagaatgctgacatcaattatgaatcagattgaacccatccctcgcaagcatt 1035
 Db 993 tggcgaagagtgcttaataacaaagcgttgctgagcagcttgatgaacccatt 1052
 QY 1036 tgaagaagctatattgctgactatccacaggtctggaagatgacgaagaaagttct 1095
 Db 1053 cgtgaaaagatgattgagacgagcttlttgaaagatgacgaagatctgggagaactt 1112
 QY 1096 ccacatgggcaagagctgctataataatcattatctctcgcccaaa 1145
 Db 1113 ctatgaagctcgccacagatagctattgtattcttctgctcgttagaa 1162
 RESULT 7
 ID AAC39544 standard: DNA; 1345 BP.

XX AAC39544;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 25021.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-030139.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 23-MAR-1999; 99US-0123548.
XX 25-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134421.
XX 18-MAY-1999; 99US-0134770.
XX 18-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 15-JUN-1999; 99US-0139452.
XX 15-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139482.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140861.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142890.
PR 08-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145616.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145819.
PR 28-JUL-1999; 99US-0145851.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0148968.
PR 17-AUG-1999; 99US-0149378.
PR 18-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158365.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158284.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 7.0%; Score 90.4; DB 21; Length 1345;
 Best Local Similarity 49.4%; Pred. No. 3,1e-14;
 Matches 380; Conservative 0; Mismatches 351; Indels 39; Gaps 4;

```

QY 415 agagatgctgcaataagcaatgctgctcttcaagagagactctcccgagga 474
DB 441 agaatcgctgttggcgccgcgcaggttcgtctcagcgaggttgcctgcgcg 500
QY 475 ggcacatgatttttccactctgttacaagtgtatgattacaggttccagcg 534
DB 501 ggccttcacttggcatctctctcttacttacttgggtgtcaggttccatgcy 560
QY 535 ttgtgattgatttggggt-----tgggcaaaagagaggtattta 579
DB 561 tgaagcgagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 620
QY 580 ctcttcaaaagagtgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 639
DB 621 catatcaaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 680

```

```

QY 640 tttaacacatttcaagatcttccgaagaagtgttttcaagtgcccaatgctct 639
DB 661 ttctcttgcttttttgaagtaagctgaggttggccgggaagcgcgaatgcttt 740
QY 700 taactgatttgt--aaagtatgatttcagacacga-----atcccttga 747
DB 741 atccttccttvgtaagaagaaacacgtgaccccaacgaagaagatgtcctaattg 800
QY 748 cttaactgacatgcaataaagacattgatttgaaggacttcggaagaaagaaatt 807
DB 801 actctatgctcaagctcttactatgctacagcaaaagaggtatcctcaggaagaagat 860
QY 808 ggaagttcaatattccattcttacccttcagcaagaagaagaaagaaagaaagaa 867
DB 861 cgatgcttcaagccttactactatgctgagcctccgaagaagaaagaaagaaagaa 920
QY 866 ggaagaggttcttcggaattttatatactggaagacttlaagcccatatgagtcgc 927
DB 921 gaagaagaggttcttcttgatcagatagcgaagaaagaaagaaagaaagaaagaa 980
QY 928 ctctctatgataagatgattaccagta-----gataccatgaaacaataaa 975
DB 981 gagatcagtgaggaagaaatgattgaccttgcaataagatccaaacgaagcctgctag 1040
QY 976 agcagagatvgggcatcaattaatatgacagtttaacgaaccatcctcgcaagcatt 1035
DB 1041 tggcgaagagtggtcctaataacaaagcgtggtcgaagcagatgtagaaccatt 1100
QY 1036 tgaagaagctatgctgctgattatccacaggttcggaagatcgcaagaaagttct 1095
DB 1101 cgggtgaanaatgattgagcagccttcttgaagaagatgcaagatcgtggaagtaact 1160
QY 1096 ccaatggaagagctgctataataatcttacttctctgcgcacaa 1145
DB 1161 ctatgaagctgcacagatagcattgttacttctgcgtctagaa 1210

```

RESULT 8

ABA05044
 ID ABA05044 standard; CDNA; 1476 BP.

XX ABA05044:

XX 22-FEB-2002 (first entry)

DE A thaliana jasmonic acid carboxyl methyltransferase coding sequence #2.

XX Jasmonic acid carboxyl methyltransferase; virus resistance;

KW Insect resistance; cold resistance; drought resistance; antifungal;

KW S-acetoxy-L-methionine; jasmonic acid carboxyl methyltransferase;

KW bacterial phytopathogen resistance; salt resistance; JMT;

XX antibacterial; antiviral; transgenic plant; ss.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers

FT 5'UTR 1..14

FT CDS /tag= a

FT 3'UTR /tag= b

FT /product= "jasmonic acid carboxyl methyltransferase"

PN MO200196549-A1

PD 20-DEC-2001

XX 05-JUN-2001; 2001MO-KR00953.

XX 13-JUN-2000; 2000KR-0032365.

PA (SCIG-) SCIGEN HARVEST CO LTD.
 XX (CHOI/) CHOI Y.
 PI Choi Y, Cheong J, Lee J, Song J, Song S, Seo H, Koo Y;
 XX WPI: 2002-090208/12.
 DR P-PSDB; AAM47309.
 XX
 PT Jasmonic acid carboxyl methyltransferase and gene encoding it useful
 PT for enhancing a resistance of plant against damages caused by
 PT phytopathogens and harmful insects, and stresses
 PS
 PS Claim 4, Page 49-51; 56pp; English.
 CC The present invention relates to the Arabidopsis thaliana jasmonic acid
 CC carboxyl methyltransferase (JMT, also known as S-adenosyl-L-methionine:
 CC jasmonic acid carboxyl methyltransferase). The sequences can be used to
 CC produce transgenic plants with increased resistance to cold, salt,
 CC drought, viruses, harmful insects and bacterial phytopathogens. The
 CC present sequence is one version of the cDNA of the invention.
 CC
 XX Sequence 1476 BP; 427 A; 285 C; 357 G; 407 T; 0 other;
 SO
 Query Match 7.0%; Score 90.4; DB 24; Length 1476;
 Best Local Similarity 49.4%; Pred. No. 3,2e-14;
 Matches 380; Conservative 0; Mismatches 351; Indels 39; Gaps 4;
 QY 415 aggacgctgcctaatagcagctcttcttcacgacacccctcccgagga 474
 DB 407 agaacgctgttctgtctgctccgctccagctgtctcctcgaagctgttctccgcg 466
 QY 475 gccatgatttttgcattctgttaacgtctcttctgttctcagctcgcg 534
 DB 467 gagccctacatctgtcagctctctcctcagttcagttcgtccagctccacg 526
 QY 535 ctctgtgactgaattgggac-----tggctcaacaaggagattta 579
 DB 527 tgaagcggagagaagagacagacataacagctgattagaacaacaggaataata 586
 QY 580 ccttcacaaagagatgtctccgctccagaaagacatttgatcaattacaaga 639
 DB 587 catatcaaaagacagctccttaagagtcacataagctctctcattccacaactga 646
 QY 640 ttttccacattccttaagattcattcgaagaagctgttctcagctcgaagctcct 699
 DB 647 tttctgttttttgggtcagctcgtgagagttgtccggagagcgcaatggttt 706
 QY 700 taacgcatttgt--aaagtagaattcgacgaacga-----atccctaga 747
 DB 707 acgcttctgttagaagatcacctgacccacaacggaagagtgctatacagga 766
 QY 748 ctacttgaatgacataaagacttattgttggagacttcctggaggaagaatt 807
 DB 767 acctcagctcaagctccttatgttcaatggccaagaggtatcactcggagagagagat 826
 QY 808 ggaatgttcaattcattcattcaaccttcagcagaagaagaatgacagttga 867
 DB 827 cgaatcttcaacgctcctcctacatgctgcgagctccggaaggttgaataatgagaga 886
 QY 868 ggaaggaagttctctgcaaatattatctcggagactttaagccactatgtctgc 927
 DB 887 gaaagaagagtcatttctcgtatcgaataggtcgaatagcttgatgttggaaggttg 946
 QY 928 ctctctctatgatgatthaccagta-----agatcccatgacaattaa 975
 DB 947 gagatcagctgagagagattgacctgcaataaagtcacaacccggaagcctagtag 1006
 QY 976 agcagagatctgacatcattatagatcagttttagaacccatcccgcaagctatt 1035
 DB 1007 tggcgaagagtgctcctatatacctaagagcgtgtgcgacgactgctagaactactt 1066
 QY 1036 tggagagacttatatgctcgaacttccaaaggcttcggaagatcgagaaggttct 1095

DB 1067 cgttgaatgtgtagacgagcttcttgaagatgcaagatcgtggagagactt 1126
 QY 1096 ccataaggcgaagcgtctcctaataatcattcattctctcgcacaaa 1145
 DB 1127 ctatgtaagctgcgacagatagcattgtattcttctcgcgtttagaa 1176
 RESULT 9
 AAC48773
 ID AAC48773 standard; DNA; 1206 BP.
 AC AAC48773;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58713.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP103405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PE
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126564.
 PR 28-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0132407.
 PR 30-APR-1999; 99US-0132484.
 PR 04-MAY-1999; 99US-0132485.
 PR 05-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134700.
 PR 18-MAY-1999; 99US-0134768.
 PR 18-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136782.
 PR 28-MAY-1999; 99US-0137528.
 PR 01-JUN-1999; 99US-0137528.
 PR 03-JUN-1999; 99US-0137502.
 PR 04-JUN-1999; 99US-0137724.
 PR 07-JUN-1999; 99US-0138094.
 PR 08-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.


```
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139466.
PR 22-JUN-1999; 99US-0139467.
PR 23-JUN-1999; 99US-0139468.
PR 23-JUN-1999; 99US-0139469.
PR 24-JUN-1999; 99US-0139470.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140824.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141288.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143543.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144006.
PR 16-JUL-1999; 99US-0144007.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147304.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149931.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150567.
PR 27-AUG-1999; 99US-0150568.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151067.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151304.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153071.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154019.
PR 22-SEP-1999; 99US-0154439.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155559.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160779.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160816.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160982.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.98; Score 50.4; DB 21; Length 1206;
Best Local Similarity 46.18; Pred. No. 0.0017;
Matches 257; Conservative 0; Mismatches 286; Indels 15; Gaps 2;

QY 32 atgaagctccagaagctcgtcatatgaatgaaggtgaagcgatatacgaacgaag 91
DB 42 atgaagctcgaagactctcagcagaaggtgcaagagacaagacgaactacgccaat 101
```

```
QY 92 aatgcatcctaactctgctcttccaaagtgaaaccttcttcgtaacaatgatacga 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 aactcaagctccagcaatgatgacgctccatcagcttcacccctcgaaggaaacct 161
QY 152 gaattgtgctgggagcaacttgcgcaatacaagtgatcaaatgctcggtatgga 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 gaaatgttcaactaactctctccgacgtccgcaacgttcaaggggttgcattg 221
QY 212 tgcctctgcgcaacaacacactttacagtgctggagcatgtgcgaagtgttaca 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 tcttcctccgagcaataacgtctccacataatcgaattcaacatacctccta 280
QY 272 gtgagcagaagaagaatgaattagaacgtccaccattcagatttctcgaagat 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 ----gaggttgcgtccgagatccgacccctccgagattcacaagcttcttcg 335
QY 332 ctcttcaaatgatcattcagtttccagtttgcgtccgaagct-----ctac 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 ctcccaagcagcagactcaacacgcttctccagcttctccacacttgcctacact 395
QY 383 cgaactcgaagaagaatgagcagcaagatagatcgttgcataaagcgaatgct 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 tgcattgagagtgcttgcgtccgacggaacgctctactctcgttgcgttctcc 455
QY 443 ggcctcttcagcagcagctcttcccgagagtgatcagcatttttgcactctgt 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 ggaacgttaccgagcagcttctccagcagacacatgacttctccactctgcctc 515
QY 503 agtctcattggtatcccaagttcccaagcgttgcgtgagatgaatgggagatgca 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 tccctgattgctcctcgaagtcgcgaagtgagcagatagaagatcgcgagctac 575
QY 563 aacaaaggagatttacc 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 aatagaggagagtttcc 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AAC33567
ID AAC33567 standard; DNA; 1211 BP.
XX
AC AAC33567;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3524.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM Protein identification; signal transduction pathway;
KM Metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 23-MAR-1999; 99US-0123548.
PR 25-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130440.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132488.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140692.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142129.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144637.
PR 20-JUL-1999; 99US-0144864.
```

```

PR 21-JUL-1999; 99US-0144814
PR 21-JUL-1999; 99US-0145086
PR 21-JUL-1999; 99US-0145088
PR 22-JUL-1999; 99US-0145085
PR 22-JUL-1999; 99US-0145087
PR 22-JUL-1999; 99US-0145089
PR 22-JUL-1999; 99US-0145192
PR 22-JUL-1999; 99US-0145145
PR 23-JUL-1999; 99US-0145224
PR 23-JUL-1999; 99US-0145276
PR 26-JUL-1999; 99US-0145913
PR 27-JUL-1999; 99US-0145918
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 03-AUG-1999; 99US-0146389
PR 04-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 05-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 06-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 09-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148365
PR 13-AUG-1999; 99US-0148684
PR 16-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149175
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149902
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150864
PR 26-AUG-1999; 99US-0150866
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066
PR 30-AUG-1999; 99US-0151080
PR 31-AUG-1999; 99US-0151303
PR 01-SEP-1999; 99US-0151336
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153700
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157153
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158025
PR 08-OCT-1999; 99US-0158325
PR 12-OCT-1999; 99US-0158363
PR 13-OCT-1999; 99US-0158393
PR 13-OCT-1999; 99US-0159294
PR 13-OCT-1999; 99US-0159295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638

```

```

PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 25-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405
PR 25-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161921
PR 28-OCT-1999; 99US-0161992
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162147
PR 29-OCT-1999; 99US-0162147

Query Match 3.8% Score 48.8; DB 21; Length 1211;
Best Local Similarity 45.9%; Prod No. 0.00477;
Matches 256; Conservative 0; Mismatches 287; Indels 15; Gaps 2;

QY 32 atggagctccagaagatcttcgataatgaaggtgaagcgagataagctacgcaag 91
DB 46 atgaagctcggagactcttcgacatgaaggtggaaggaagacagacatcgcaat 105
QY 92 aatgatccatcatctgctcttcgcaaggtgaaccttctcttgacatgacatga 151
DB 106 aactctcaagctcagacatgacatgacatgacatgacatgacatgacatgacat 165
QY 152 gaattgttcggcgcaacttcgcaacacacacacacacacacacacacacacacac 211
DB 166 gaaatgttcaccccaacccctccgacgacgacgacgacgacgacgacgacgacgac 225
QY 212 tccgcttcgacacacacacacacacacacacacacacacacacacacacacac 271
DB 226 tcttccttcgacacacacacacacacacacacacacacacacacacacacacac 284
QY 272 gtcttcgcaaggaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 331
DB 285 -----gaagttcgtatcgtcggaaatcgacccctcggaatgaatgaatgaatgaat 339
QY 332 ctttcacaaatgattcaatcgtttcaagtgtgcgcaagctt-----ctac 382
DB 340 cttcccaagcagcacttcaacacacacacacacacacacacacacacacacacacac 399
QY 383 cgaacatcgagaagaagaatgaagcaagatagatcgttgccataaagaagcaatgact 442
DB 400 tgcagagagaggtctcttcgacgacgacgacgacgacgacgacgacgacgacgac 459
QY 443 ggccttcctacgacgacactctcccgagagatgcacatgacatgacatgacatgac 502
DB 460 ggaatgtttacgacgacacttttcacgacgacacacacacacacacacacacacac 519
QY 503 agtgcattgttatcaggttccacgacgacgacgacgacgacgacgacgacgacgac 562
DB 520 tccctgcattgcctcccaaggtcgcgaagaagtcgaaggaagatcgcgcgacac 579
QY 563 aacaaaggagatttac 580
DB 580 aatagagggagatttcc 597

RESULT 11
AA07727/C
ID AA07727 standard; DNA; 198 BP.
XX AA07727;
AC
XX
DT 23-OCT-2001 (first entry)

```


Db 525 aagtgacgagatagagatcgccgcgcacacatagaggagagtttc 573
RESULT 13
ID AAC36551 standard; DNA; 1261 BP.
AC AAC36551;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14222.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN BP103405-A2.
PD
PD 06-SEP-2000.
DF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132483.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143634.
PR 13-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148668.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

```

PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149902
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151085
PR 27-AUG-1999; 99US-0151086
PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157665
PR 07-OCT-1999; 99US-0158028
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158359
PR 13-OCT-1999; 99US-0159283
PR 13-OCT-1999; 99US-0159284
PR 14-OCT-1999; 99US-0159285
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 25-OCT-1999; 99US-0160989
PR 25-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405
PR 26-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161353
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161922
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

```

Query Match 3.3%; Score 42.6; DB 21; Length 1261;
 Best Local Similarity 53.3%; Pred. No. 0.22;
 Matches 90; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

OY 412 gatagatcgttcgcaataagcgcacatgctgcgtctcttcttaagcagacatcccca 471
DB 407 gaacgcctcctactcgtcgtcgtgttcccgatcgcttcttaacggtagcttctccagc 466
OY 472 ggaagtcacgacttttttgacactcgttaacagtgatcgttatactcaggtccag 531
DB 467 gagaacatcgtacttctccacacgcgcctcctcctcgtcattgagctccacaggtccaga 526
OY 532 cggcttgagatgattgggagatcgtgagcaacaaggagattattac 580

```

DB 527 aagttgacgagatagagatcgcgcgcgtacataagagggaggttttc 575

RESULT 14

ID AAX60802 standard; DNA; 3637 BP.

AAX60802;

09-AUG-1999 (first entry)

Human secreted protein encoding DNA (clone fj283-11).

Secreted protein; kidney; lung; brain; blood; testis; bone marrow;

nutritional activity; cytokine; cell proliferation; immune stimulation;

hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;

anti-inflammatory; tumour invasion; ss.

Homo sapiens.

W09926961-A1.

03-JUN-1999.

24-NOV-1998; 98MO-US25149.

23-NOV-1998; 98US-0197886.

26-NOV-1997; 97US-0066804.

(GENEX) GENETICS INST INC.

Agostino MJ, Clark HF, Collins-Racie LA, Evans C;

Rechtel K, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Steininger RJ, Treacy M, Wong GG;

WPI: 1999-357809/30.

P-PSDB: AAY17220.

New polynucleotides encoding secreted proteins

Disclosure: Page 113-114; 133pp; English.

The invention relates to secreted proteins (AAV17219-228) encoded by

polynucleotides obtained from human fetal kidney, adult lung, adult

kidney, adult brain, adult blood, adult testes, and fetal brain and

murine adult bone marrow cDNA libraries. The secreted protein nucleic

acid sequences (X6801-811) correspond to clones D0306-7, 9J283-6,

IK317-3, K213-2x, na16-1, n193-20, nP164-1, pe204-1, yal-1 and yd-1.

(all clones are deposited as ATCC 98599); The PNs and proteins are

predicted to have biological activities which would make them suitable

for treating, preventing or ameliorating medical conditions in humans and

animals, although no supporting data is given. Suggested activities

include nutritional activity, cytokine and cell

proliferation/differentiation activity, immune stimulating (e.g. as

vaccines) or suppressing activity, hematopoiesis regulating activity,

cell tissue growth activity, activin/inhibin activity,

chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,

receptor/ligand activity, anti-inflammatory activity, cadherin/tumour

invasion suppressor activity, and tumour inhibition activity. The PNs are

also stated to be useful for gene therapy.

Sequence 3637 BP; 1132 A; 598 C; 837 G; 1069 T; 1 other;

Query Match

Best Local Similarity 3.2%; Score 41.4; DB 20; Length 3637;

Matches 72; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 1176 tttttagttgttttttgccggtgggggtccttcggggtatcgttttgatccgtaa 1235

DB 3487 ttttggtagctttgttttgccggtgggggtccttcggggtatcgttttgatccgtaa 3546

OY 1236 taaagtgatgctgcagaataagatatttagtaacatttttcacataaaaaaaaaa 1295
 DB 3547 taatagtaacataaaggttaacatttttaagtgaataaaaaaaaaa 3606
 OY 1296 aaa 1298
 DB 3607 aaa 3609

RESULT 15
 ID AAS59208
 AC AAS59208 standard; cDNA; 3637 BP.
 XX AAS59208;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human cDNA encoding a secreted protein f3283_11.
 XX
 KW Human; secreted protein; ss; antiinflammatory; immunosuppressive;
 KW neurotrophic; neuroprotective; antiarthritic; antimicrobial; vulnery;
 KW cytosolic; antidiabetic; virucide; antifertility; anticonvulsant;
 KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KW antihemagic; antitumor; antilucer; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement.
 XX
 OS Homo sapiens.
 XX
 PN NC0200175068-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 22-MAR-2001; 2001MO-US09369.
 XX
 PR 30-MAR-2000; 2000US-0539330.
 PR 04-DEC-2000; 2000US-0729674.
 XX
 PA (GENM) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie E, Collins-racle LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GC;
 PI Clark H, Fechtel K, Werberg D;
 XX
 DR WPI; 2001-639363/73.
 DR P-PSDB; NA038990.
 XX
 PT Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke).
 XX
 PS Disclosure; Page 459-460; 619pp; English.
 XX
 CC The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD). In the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair.
 CC In the treatment of burns, incisions and ulcers, as well as in treatment
 CC of periodontal disease, osteoporosis or osteoarthritis, mediated by

CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 CC infection of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activity or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibins to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC female mammals and nucleic acids are also useful as food supplements. The
 CC present sequence encodes a secreted protein of the invention.
 XX
 SQ Sequence 3637 BP; 1132 A; 598 C; 837 G; 1069 T; 1 other;

Query Match 3.2%; Score 41.4; DB 22; Length 3637;
 Best Local Similarity 58.5%; Pred. No. 0.72;
 Matches 72; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 OY 1176 ttttaagtgcttttgcgcgttggggcttcgcggtatgctgttgcgtatgaa 1235
 DB 3487 ttttggctgcttttgcgttgcggtatgctgttgcgtatgaa 3546
 OY 1236 taaagtgatgctgcagaataagatatttagtaacatttttcacataaaaaaaaaa 1295
 DB 3547 taatagtaacataaaggttaacatttttaagtgaataaaaaaaaaa 3606
 OY 1296 aaa 1298
 DB 3607 aaa 3609

Search completed: July 27, 2002, 04:45:04
 Job time: 4660 sec

630-0101, Japan (E-mail: sano@bmailgate.aist-nara.ac.jp,
 URL: http://gicw3.aist-nara.ac.jp/sano.html,
 Tel: 81-743-72-5652, Fax: 81-743-72-5659)

FEATURES
 source 1..1316
 Location/Qualifiers

gene
 /organism="Coffea arabica"
 /cultivar="catuira"
 /db_xref="taxon:13443"
 45..1163
 /gene="CamTL3"
 45..1163
 /gene="CamTL3"
 /codon_start=1
 /product="theobromine synthase"
 /protein_id="BAB39215.1"
 /db_xref="GI:13365751"
 /translation="MELOEVLNMGEGEDTSYKNSAYQVLAVKRPVLEQCVRELL
 RANLPNKKCIKIVADIGCASGPNLLIVRDVOSIDKQEKNELEKPTIQIFLNDL
 FNDNSVYFKLLPSEYFKLEKENGKRIQSCIGMGSFYSRLPFESMHPLHSCYCL
 OHLISOVPSGLVTEIGISTNKGSISYKSKASRLPYOKAYLDQFSPKDFPLRHSLELS
 HSRMLTICIKGYELDRNADIDLEMAINDVLEEKLEKDSFNLPYIPSAEYK
 CIVEESEFEILYETFEKYLVDYGSIDDEHKAREVASVAVYEPILASHGEAL
 PDIHPRKHAATVLEPGKGFYNNLTLSLAKKPKSDV"

CDS
 386 a 250 c 292 g 388 t
 BASF COUNT
 ORIGIN

Query Match 78.3% Score 1016.4; DB 8; Length 1316;
 Best Local Similarity 89.0%; Pred. No. 5,96-252;
 Matches 1162; Conservative 0; Mismatches 96; Indels 48; Gaps 4;

19 gtcctcatatgaatgagctccaaagaagctctgcatatgaatgaagtgagcatatc 78
 32 gtccgcgcatatgagatgagctccaaagaagctccgagatgagatgagcatatc 91
 79 aagctacgccaagaatgcatcctacaat--ctgctctctgccaagtgagaccttctc 135
 92 aagctacgccaagaatgcatcctacaat--ctgctctctgccaagtgagaccttctc 151
 136 tgaagaatgcatcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 195
 152 tgaagaatgcatcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 211
 196 aattcgagatctgagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 255
 212 aattcgagatctgagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 271
 256 gcaagaatgcatcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 315
 272 ccaagaatgcatcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 331
 316 gattttctgagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 375
 332 gattttctgagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 391
 376 ctctacgccaactgagatcgaagaatgagcgaagatagatcgttcccaataagcgc 435
 392 ctctacgccaactgagatcgaagaatgagcgaagatagatcgttcccaataagcgc 451
 436 aattcgatctgtcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 495
 452 aattcgatctgtcgaagatctgtcggggccaactgtcccaacatcaacaagtgcatlaa 511
 496 tggtaagtgatcattgattctcgaagtgcatcgaagtgcatcgaagtgcatlaa 555
 512 tggtaagtgatcattgattctcgaagtgcatcgaagtgcatcgaagtgcatlaa 571
 556 tggtaagtgatcattgattctcgaagtgcatcgaagtgcatcgaagtgcatlaa 615
 572 cagtgacgaacaaagagcatttactcttccaaagatctgtccgcgcgtccgaagtc 631
 616 aatttgatcattcgaagatcttcccaacatcttcaagatctcgaagatct 675

632 ATATTGCATATTAGCAAGATTTTACACATTCTAGATTCATTTCGAGAGT 691
 676 gtcttcagggccgaagatcctcctccgcatcttgtaagatgagatctcgaagc 735
 692 GTTTCACATGCGCCATGCTCTCTTACTTGATTTGTAAGAGTGAATTAGCGCCG 751
 736 gaatccctagatcattgacatgacataaagacttgattgttgagggatcttgg 795
 752 GAATGCGATGACTTACTTGTAGATGACATTAACGACTTGATTGAGGACATCGA 811
 796 ggaagaataattgagatcttcaatcattcctcaccctcagcaggaagataa 835
 812 GGAAGAAATTTGATGATTCATCTCCAGATTAATACCTGACGAGAAAGATTA 871
 856 gtgatatgttgaaggaagatcttctcgaaatcttataatcggagactttaaggcca 915
 872 GTGATATGTTGAGAGGAGGATCTTTTGAATTTTATATCCTGAGACTTTTAAAGTCT 931
 916 ttatgatgcctctctcattgattgattaccgaagatcccatgaacaaatlaa 975
 932 TTACGATGCTGCGCTTCTATTTGA-----CGATGAACATATTAA 970
 976 agcagatgctgcatcattatagatcagatcttaccgaacccatcctcgaagcatct 1035
 971 AGCAGATGCTGATGATCTCCGTTAGAGATTTAGACACCATCTCGAAGCATTT 1030
 1036 tggagaagctatgactgactgactcagagatcgaagatcgaagatcttct 1095
 1031 TGGAGAGCTATATATCTGACATATTCACAGGTTGCGAGATGACACCAAGGTTCT 1090
 1096 ccaatgaggaagcgtgcatlaaataatcattcattcctcgcgaagcgaagaa 1155
 1091 CCCCTTGGGCAAGGCTTCTATTAATTTATTCATTCATTTCTCGCCAAAACCGAGAA 1150
 1156 gtccagcgtgtaaaagatctgtcttcttcttcttcttcttcttcttcttcttcttct 1207
 1151 GTCCAGAGCTGTAAGATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1210
 1208 ttcgggattctg-----tcgtttgattcgttaataaagtgatgca 1251
 1211 TTCGGGATTGCTCTTTATTAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1270
 1252 gataaagatctttagacatattcttataaataaataaataaataaataaataaataa 1297
 1271 GAATTAATATTGACATATTAATTTTCAAAAAAATAAATAAATAAATAAATAAATAA 1316

RESULT 3
 AB048792 1304 bp mRNA linear PLN 13-MAR-2001
 LOCUS
 DEFINITION Coffea arabica CamTL2 mRNA for theobromine synthase, complete cds.
 ACCESSION AB048792
 VERSION AB048792.1 GI:13365748
 KEYWORDS
 SOURCE
 ORGANISM
 Coffea arabica (cultivar:catuira) cDNA to mRNA.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Gentianales; Rubiaceae; Ixordiales;
 Coffeaceae; Coffea.
 1 (sites)
 REFERENCES
 1. Ogawa, M., Heral, Y., Koizumi, N., Kusano, T. and Sano, H.
 7-Methylxanthine methyltransferase of coffee plants. gene isolation
 and enzymatic properties
 J Biol Chem. 276 (11), 8213-8218 (2001)
 2 (bases, 1 to 1304)
 Sano, H., Ogawa, M., Kusano, T. and Koizumi, N.
 Direct Submission
 Submitted (11-SEP-2000) Hiroshi Sano, Nara Institute of science and
 technology, Plant molecular breeding, 8916-5 Takayama, Ikoma, Nara
 630-0101, Japan (E-mail: sano@bmailgate.aist-nara.ac.jp,
 URL: http://gicw3.aist-nara.ac.jp/sano/sano.html,

```

FEATURES
    Tel:81-743-72-5652, Fax:81-743-72-5659)
    Location/Qualifiers
        source
            1..1304
                /organism="Coffea arabica"
                /cultivar="catuia"
                /db_xref="taxon:13443"
        gene
            44..1201
                /gene="Camt12"
                /gene="Camt12"
                /codon_start=1
                /product="theobromine synthase"
                /protein_id="BAB39214.1"
                /db_xref="GI:13365749"
        CDS
            44..1201
                /translation="MELQVLAHNGSGDASVAKNSNOLYLAKYKPVLEQVCELL
                RANLPINIKIRVADIGAGSGPNLLTVRPIIOSIDKIRPOMKRELPERPIQVPTDL
                FQNDPNSVEMLLPSFRIKREKNGKISGLTILAMPKSPFRLPEPESNAHFIASSYSL
                OFLSQVSGIYVEIGITANKRSIYSSTASPPQAKYADQKTFQFTLAISSYSL
                BGRMLNLCICKDEPDGPNMDLEAMINDLVYEHSEHLEKDSFENPFIITASVEELK
                CIVRPSFEIILTEFKRTDQSFSDIDCVSRHSFSEYDEHARAHAVASLNSVY
                ERLANFGAIIIPDIFHRFATNAKVIIRLGSFTNNLLIISLAKPKESDI"
BASE COUNT      388 a      261 c      285 g      370 t
ORIGIN
Query Match      75.08; Score 973; DB 8; Length 1304;
Best Local Similarity 89.2%; Pred. No. 9.8e-241;
Matches 1080; Conservative 0; Mismatches 110; Indels 21; Gaps 2;

QY 19 gttcctgatatgaatgagagcccaagaagtcctgatatgaatgagagtgagagcgatc 78
DB 31 GTCCTGATATGAAAGAGAGCCCAAGAGTCTGCTGATGAAAGAGAGCGGAGGAGATGC 90
QY 79 aagtaagcaagaatgacatcctcaaat---ctggccttggccaaggtgaaccttctc 135
DB 91 AAGCTAGCGCAAGAAATTCATCTTCATCACTGCTTCTGCCAAGGTGAACCTGTCCT 150
QY 136 tgaacaatgatatgagatgttggcgagcaactggccaataataaagtgatgaa 195
DB 151 TGAACATGCGTAGAGGAAATGTTGGGGCCAACTGGCCCAATCAAAATGATCATTTA 210
QY 196 agttggagattggagatgccttcctgcaccaaacacatttaacagtgaggacattgt 235
DB 211 AGTTGCGGATTGGATGCGCTTCGACCAACACACTTTTAAAGTTGCGGACATTGT 270
QY 236 gcaaaagtatgacaaagtctggccaagaagaagaatgaattgaacgctccacattca 315
DB 271 ACAAAGATTTGACAAAGTTAGCAGAAATGAAATGAATTAAGAACGTCGCCACCATTTCA 330
QY 316 gatttctgaaatgactcttcccaaaatgattcaatcgatttcaagtgcggcaag 375
DB 331 GATTTCGACTGATCTTTTCCAAATGATTTCAATCGGTTTTCATGCTGCGCAAG 390
QY 376 ctctacccgcaaacctcgagaaagaatgaagcagatagatgctgcttaataaagcgc 435
DB 391 TTCTACCGCAAACTTGAAAGAAATGACGCAAGATGATGCTGCTCATATAGCCGC 450
QY 436 aatgcctgctcttctcagcgagacacttcccgagagatccatgcatcttttgaactc 495
DB 451 AATGCTGCTCTTTCACGCGCAACCTCTCCGAGAGAGCAATCATTTTATTCAC 510
QY 496 ttgttcaagtgtcaattgggtatcctcaaggtccagcggttggatgattgaggggat 555
DB 511 TTCTACGCTCAATTTTATTCAGGTTCCAGCGGTTTGTATGATGATGAGGAGAT 570
QY 556 tggatgcaaaaaggagatattactcttccaagaagtgcctgcgcgcgcgcagaagc 615
DB 571 CACTGCAAAAGAGAGCATTTACTCTTCCAAAGATCCCTCCCGCTCCGAGAGGC 630
QY 616 atatttgcattcaatgcaaaagatttaccacatttctaaggattcctgaaagatt 675
DB 631 ATATTGATCAATTTAGCAAAATTTTACCAATTTTAAAGATTGCTCGGAGAGTT 690

QY 676 gtttcaagtgccgaatgctccttaccctgcttcttgaagaatgaaatgcaagcagacc 735
DB 691 GCTTTACCGCGCCCAATGCTCCTTCTGCTTTCGAAAGAGATGAAATTCGAGGCC 750
QY 736 gaatccctcagacttactgacatggcacaataaagacttgaattgttggagagcttcgga 795
DB 751 GAATACCATGATGACTTACTGATGATGATGATGATGATGATGATGATGATGATGAT 810
QY 796 ggaagaagaatgagatgattcaatattccattcttacccttacccttacccttaccct 855
DB 811 GGAAGAAATTTGACAGTTTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCAATG 870
QY 856 gtccatagtgagagagagagagagagagagagagagagagagagagagagagagag 915
DB 871 GTGATAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 930
QY 916 ttatgagtgctcctctctatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 962
DB 931 TTATGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 990
QY 963 ----atgacaatlaagcagagatgagatgagatgagatgagatgagatgagatgagat 1017
DB 991 CACGATGAAACATGCTGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1050
QY 1018 catcctcgaaatcatttggagaagcattatgctgacttacccttacccttacccttacc 1077
DB 1051 CATCTCGCAATCATTTTGAAGAAACCTTATGATGATGATGATGATGATGATGATGAT 1110
QY 1078 gcatgagagagagagagagagagagagagagagagagagagagagagagagagag 1137
DB 1111 GAATGAGCAAAAGGTTATGCTGCTGCGCAAGGCTTCTATATATCTTATCATTTCTCT 1170
QY 1138 cgcgaagaagcagagagagagagagagagagagagagagagagagagagagagagag 1197
DB 1171 TGCCAAAACCAAGAGATGACATATTAAGACTGTTATGATGATGATGATGATGAT 1230
QY 1198 ttgggggctt 1208
DB 1231 TGCTTTT 1241

RESULT 4
AB039725 1360 bp mRNA linear PLN 13-MAR-2001
LOCUS Coffea arabica CamT11 mRNA for caffeine synthase, complete cds.
DEFINITION AB039725
ACCESSION AB039725.1 GI:13365693
VERSION 1
KEYWORDS caffeine synthase.
SOURCE Coffea arabica callus CDNA to mRNA.
ORGANISM Coffea arabica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Rubiaceae; Ixoroideae;
Coffeaceae; Coffea.
REFERENCE 1 (bases 1 to 1360)
Ogawa,M., Heral,Y., Koizumi,N., Kusano,T. and Sano,H.
7-Methylxanthine methyltransferase of coffee plants: gene isolation
and enzymatic properties
J. Biol. Chem. 276 (11), 8213-8218 (2001)
JOURNAL MEDLINE 21269383
AUTHORS Sano,H., Kusano,T. and Ogawa,M.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) Hiroshi Sano, Nara Institute of Science and
Technology, Plant Molecular Breeding; 8916-5 Takayama, Ikoma, Nara
630-0101, Japan. (E-mail: sano@bs.ais-t-nara.ac.jp,
Tel:+81-743-72-5650, Fax:+81-743-72-5659)
FEATURES
    source
        1..1360
            /organism="Coffea arabica"
            /db_xref="taxon:13443"
            /tissue="callus"
        14..1171

```

```

/ gene="Camt1.1"
/ 14. 1171
/ gene="Camt1.1"
/ product="caffeine synthase"
/ protein_id="BAB3213.1"
/ db_xref="GI:1336594"
/ translation="MELOEVLHNMGEENGEASVAKNSSPNOLVLAKEVPLEOCVRELL
RAULPINKICIVADGASPNITLTVWDVQSDIKVKQEMKNELEPTIOVFLTDL
FQDNFNSVFMLEPSEYRLKEKENGRIKISGLIAAPGSGFQRLPEESMFLHSYSL
QFDSQVSGVLVEIGITPANKRSIYSKASBPVQAYADQTXPTFLMRSBELS
RGMILTICIKDECDGPNMDLEMAINDLVAERLDEKLSFNVPITYTASVEEVK
CMVEEGSEFELYOTFKRLDAGFSDIDCCVSHSPVSDENARAHVASISIVY
EPLIASHFEALIPDIFRFRATNAAVVIRLGGFNNI1SLAKPEKSDI"
BASE COUNT      397 a      266 c      296 g      401 t
ORIGIN
Query Match      74.5% Score 966.6; DB 8; Length 1360;
Best local similarity 88.9% Pred. No. 4,4e-239;
Matches 1076; Conservative 0; Mismatches 114; Indels 21; Gaps 2;
QY 19 gtcgcgcataatgagatgcgaagaagcttcgacataaagagtgagagggagac 78
DB 1 gtcgcgcataatgagatgcgaagaagcttcgacataaagagtgagagggagac 60
QY 79 aagctacgcgaagaatgcatacaat---ctggctcttgccaaagtgaaacttctc 135
DB 61 AAGCTACGCCAAGAAATTCATCAATCACTGGTCTCGCCAAAGTGAAACCTGCTCT 120
QY 136 tgaacaatgacataagaaatgtgtgagggcaacttgccacaatcaaaagtacataa 195
DB 121 TGAACAATGCGTAGGGGAATGTTCGGGGCCAACTGGCCAAACATCAACATGCAATTA 180
QY 196 agtgcgagattggagatgcgtctctgacacaacacactttaaagtgaggacatgt 255
DB 181 AGTTCGAGATTGGGAGGCCCTCCGACCAACACACTTTTAACTGGGAGACACTGT 240
QY 256 gcaaaatltcaacaagtgtgccaagaagaagaagaatgaatgaacgcgccacattca 315
DB 241 ACAAAATATTGCAAAAGTAAAGCAAGAAATGAAGAAATTAAGAAAGTCCACACATTCA 300
QY 316 gatttctgaaatgctcttccaaaatgattcaatgattcaatgattcaatgattca 375
DB 301 GGTTCCTGACGTACTCTTTCCAAATGATTTCAATTCGGTTTCATGCTCTGCCAAG 360
QY 376 ctctacacgcaaaactcgaaaagaagaagaagaagaatgagacgctgctcctataagc 435
DB 361 CTCTACCGCCAAACTGAGAAAGAAATGAGAGCAAAATGAGATCGTCCATAAGCCGC 420
QY 436 aagcctgagctctcttcaaggagacacttcccccagagagatcagatcttctgactc 495
DB 421 AATGCTGAGCTCTTCCACGCGAGACTCTTCCCGAGAGATCCATGCAATTTTACACTC 480
QY 496 ctgttacaagtltcaatgttactcaggttcccccagcggttggatgattgagat 555
DB 481 TTCTTACAGTCTGAGTTTATCCAGAGTTCGCCGCTTGGTATCGAATTTGGGAT 540
QY 556 tggtagcaacaagaagagatcttacccttcccaagaagatgtctcccccgcacagaagc 615
DB 541 CACTCGGAACAAAGAGGACATTACTCTCCAAACAGATCTCCGCCGCTCCAGAAAGC 600
QY 616 atatttgagtaaatgagaagaatttccacatttcaagaatcattccagaagaatt 675
DB 601 ATATTGAGTCAATTTAGGAAAGATTTTACCAATTTTAAGAGATGCGTCCGAGAAATT 660
QY 676 gttttagagtgagcgaatgtctcttaccgcatctgttaagatgagatgagaaac 735
DB 661 GCTTTCAGCTGCGCAAGGCTCTTACTTGCATTGTGAAAGAGATGAATGAGAGCGCC 720
QY 736 gaatcccttagactactgacatggaataaagaacttgattgtgagagcttcgga 795
DB 721 GAATACCAATGACTTACTTGAATGAGCAATTAACCACTTGCTGAGGAGCTGTGG- 780

```

```

QY 796 ggaagaagaattgataatgattcaattcattccttataccttcacagaagaagataa 855
DB 781 GGAAGAAAATTTGACAGATTTCAATGTTCCATCTATACGCTTCAGTAAGAAATAAA 840
QY 856 gtgcattgttgagaggaaggttctctcgaaatttatactcggagactttaaagccca 915
DB 841 GTGCATGTTGAGAGGAAGGTTCTTTGAAATTTATCTTCGACGACTTTTAAGCTCCG 900
QY 916 ttatgagtcgctctctctatgatgatgattaccagaagaatcc-----962
DB 901 TTATGATGCGGCTTCTCATGTATGATGATTCGAAGTATGATCCATTCCTCCAGTARA 960
QY 963 -----atgaacaataaagaagatgtgagcatcatcaatgaatgaatgattacgaac 1017
DB 961 CAGCATGACATGATGATGAGCAGCGCATGTCATCATTAATTAATGATGATTAACGACC 1020
QY 1018 catctcgaagtcatttggagaaagcattatgcctgactcattccacagagctcgaa 1077
DB 1021 CATCTAGCAAGCATTTTGGAGAAAGCATTAATACCTGACATATTCACAGTTTGCAGC 1080
QY 1078 gcatgcaagaaggttctcacaatgaggaagctgctataatctatctatctctct 1137
DB 1081 GATTCACGCAAAAGTTATCCGCTTGCGAAAGCTCTATATATCTATCATTTCTCT 1140
QY 1138 ggcacaagaagcagaagatcagagcggttaaagttgttttagtggttttgccg 1197
DB 1141 TCCCAAAAACGAGAGATGACATATTAAGCTTGTATTAGTTGGTTTGTTTA 1200
QY 1198 ttggggggtcct 1208
DB 1201 TGGGTTGTTT 1211

RESULT 5
AB049752 1437 bp mRNA linear PLN 16-MAR-2001
LOCUS
DEFINITION
  Atropa belladonna ABSAMT1 mRNA for
  S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase,
  complete cds.
ACCESSION
  AB049752
VERSION
  AB049752.1 GI:13366160
KEYWORDS
  SOURCE
  Atropa belladonna (strain:M8) root cDNA to mRNA, clone:lambda
  ABSAMT1.
ORGANISM
  Atropa belladonna
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Asteridae; eusterids I; Solanales; Solanaceae; Atropa.
REFERENCE
  Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and
  Yamakawa,T.
  Cloning and expression of salicylic acid inducible and active
  S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase
  in transformed root culture of Atropa belladonna
  Unpublished
  2 (pages 1 to 1437)
JOURNAL
  REFERENCE
  Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and
  Yamakawa,T.
  Direct Submission
  Submitted (10-OCT-2000) Takashi Yamakawa, The University of Tokyo,
  Department of Global Agricultural Sciences, 1-1, Yayoi 1-chome,
  Bunkyo-ku, Tokyo 113-8657, Japan
  (E-mail:ayama@mail-ec.u-tokyo.ac.jp, Tel:81-3-5841-7515,
  Fax:81-3-5841-5304)
FEATURES
  source
  1. 1437
  /organism="Atropa belladonna"
  /strain="M8"
  /db_xref="taxon:33113"
  /clone="lambda ABSAMT1"
  /tissue="root"
  62..1135
  gene

```

CDS	/gene="AbsAMT1" 62.1135 /gene="AbsAMT1" /codon_start=1 /product="S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase" /protein_id="BAB39396.1" /db_xref="GI:13366161" /translation="MKVEVILHNGSGNDISVANNISLVORXIIITKPIREFOAISIDY CSFFPELTIADLGSSGANTFLVSVLIVKPERKINLOSANLHPDIDGNF NTIFOSLKFODLROKGEFPGCFSPGSPYRIIPRSELIPHSISLIMWLSQ VPLIRKKNENIYIATSPSPVITAYIKOIKRFSNPLKREELKMGKAVLFLUR ESPDPSKCCITIMELSMALNELVGLIEERKVDSPNIPITSPREVAIVKEKG SPINRLEATRHANVSENGINGENAKRAVPELIVSOFDQKLNLVFOYEBI ISDCISREKTEFINIVLSLTKR"									
BASE COUNT	488 a	225 c	290 g	434 t						
ORIGIN										
Query Match	13.6%	Score 176.4	DB 8	Length 1437						
Best Local Similarity	55.7%	Pred. No. 7.1e-35								
Matches 508	Conservative 0	Mismatches 356	Indels 48	Gaps 7						
QY	31	aatgagctccaagaagctcgtcatatgaatgaagtgagcgcatcaagctcagccaa	90							
DB	61	AATGAGGCTTGTGACAGTCTTCACATGATGAGAAATGTGACATTCATATGCAAA	120							
QY	91	gaatgc---atccataatctgcttcgcaagtgaaaccttcttaaatgaatgat	147							
DB	121	CATTCCTCTGTTGACAGAAAGTAATCTCATACCAAGCCAAATACAGAACGCCAT	180							
QY	148	acgagaatctgtgagggcgaacttgcaccaacatcaagaatgcatatgaatgcgatt	207							
DB	181	AGGTATCTCTACGACAGCTCTTCCAGAAA-----CCTTAAGCATTCCTATT	231							
QY	208	ggatgagctcttgaccacaacacactttaaagtcgagacatctgcaagatata	267							
DB	232	GGGTGTTCTTCCGAGACGAACTTTTGGTGTATCAACATTTTAAATCCTCTCA	291							
QY	268	caaatctgccaagaagaagaatgaatgaagtcaccacatlaagttttcttga	327							
DB	292	AAAAAGAAAGAAATACATTAATCTTCAGTACGGGGA-----AATCTTTTCATCTCA	345							
QY	328	tgatcttccaaatgaattcaatccggttttaagctgtgccaagctcttaacgaa	387							
DB	346	TGATCTCTCTGCAATGATTTTAAACATTTTTCAGTCATTTGGGGAATTTCAACAA	405							
QY	388	actgagaagaagaatgagcaagaatagatcgtgctataagcgcaatcctgctc	447							
DB	406	TTTGAAAGAAAGTAATGAGAAAGATTTGTCATGTTTATGTAAGTGGCTGGTTC	465							
QY	448	ttctaagcgagactctcccgagagatcagatcttttgcaactctgttacagtg	507							
DB	466	ATTATTAATCTAGACTTTTCCCTCAGAGAGTTTGATTTTTCACCTCCAGTTACAGTCT	525							
QY	508	tcattggtatctcaggttcccgaggttggatgaattggggaattggtgcaacaa	567							
DB	526	CATGCGCTATCTCAACTTCC-----TGATTTAATTTGAAAA-----GAAACA	567							
QY	568	agggagatcttaactctccaagaagatgctgcgcgcgctccagaagaagcatattgata	627							
DB	568	GGGAAACTTATTCATTCAGAGTACAGTCCACAGTGTATTAAACACATACAGCAAGA	627							
QY	628	attacagaagaatttaccacatttctaaggtatcagaagaagtgatttccagtg	687							
DB	628	ATACGAAAAAATTTTTCATATTTCTCAAAATACCGTTCCGAGGAATATATGAAGCTG	687							
QY	688	cgaatgctcttaccctcgcatctgtaagtgatgaa-----ttcagaagaacgaatccc	742							
DB	688	GAAAAATGTTTAAACATTTTAAAGAAAGTATGATCTCTTTCGCAAAAGATGCTG	747							
QY	743	ctagactacttga-----catgcaataaagacttgatctgtgaaggacttga	795							
DB	743	CTAGACTACTTGA-----CATGCAATAAAGACTTGATCTGTGAAGGACTTGA	795							

DB	748	CTATATTGGACCTTTATTCATGAGCCCTTAATGATGTTTGAAGGATGTATAGA	807																					
QY	796	ggaagaagaattgagatgattcaattccattcttaacaccttcagagagaagtaa	855																					
DB	808	AGAAGAAAGTGAATGATTCATCATATTCCTCAATACCAATCAACAGAAAGTAA	867																					
QY	856	gtgagatgtagaagaagtgcttcgaaatttatcttgagactttaagccca	915																					
DB	868	GTAATAGTTGAAAGAAAGTTCATTCATTAATCAATGATGAAGCTTACAGAGTCA	927																					
QY	916	ttatgagctgc 927																						
DB	928	TTGGAATGTTTC 939																						
RESULT	6																							
SFL308570	1152 bp	mRNA	linear	PLN 01-MAR-2001																				
LOCUS	Stephanotis floribunda	mRNA for S-adenosyl-L-methionine:salicylic																						
DEFINITION	acid carboxyl methyltransferase (samt gene)																							
ACCESSION	AJ308570																							
VERSION	AJ308570.1	GI:13235640																						
KEYWORDS	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase; samt gene.																							
SOURCE	Stephanotis floribunda																							
ORGANISM	Stephanotis floribunda																							
REFERENCE	1 (bases 1 to 1152) Potl,M.B., Pichersky,E. and Piechulla,B. Circadian oscillation of methyl salicylate emission, samt enzyme activity, and samt mRNA in flowers of Stephanotis floribunda Unpublished																							
REFERENCE	2 (bases 1 to 1152) Piechulla,B. Direct Submission Submitted (28-FEB-2001) Piechulla B., Department of Molecular Physiology and Biotechnology, University of Rostock, Gertudenstr. 11a, 18051 Rostock, GERMANY																							
FEATURES	1..1152 location/Qualifiers																							
source	1..1152 /organism="Stephanotis floribunda" /db_xref="taxon:85938"																							
gene	29..1129 /gene="samt" 29..1129 /gene="samt" /codon_start=1 /evidence="experimental /product="S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase" /protein_id="CAC33768.1"																							
CDS	1..1152 /translation="MEVEVILHNGSGTGDASYASNSLLOKKVILLTKPIREAFITELY TLFPRKSIDMCGSSGPNTPFLAVSELIKNKKRTSLGSEPEYOIHLNDPSNDF NTIFSLSPFKSPKSGMGSGPCHPFTGVPSPFGRFLPKSLHFHSSYSIMWLSQ VPLDEEVKNQIVLSTSPISVITAYIKOIKRFSNPLKREELKMGKAVLFLUR KEDHSKSGSYALELIALNELVSEGIIEEDIDCENNVQYPPSPAEKKEVVEEG SSITRELATITHTLAVDHDHVTGHHHAFKDGYSLSNCVAVPELIVVHRREALMD EYFHHREPLINCMETKEXIEFINVYSMKRRY"																							
BASE COUNT	345 a	236 c	268 g	303 t																				
ORIGIN																								
Query Match	13.0%	Score 169	DB 8	Length 1152																				
Best Local Similarity	52.9%	Pred. No. 5.8e-33																						
Matches 557	Conservative 0	Mismatches 445	Indels 51	Gaps 7																				
QY	32	atggaagctccaagaagctcgtcatatgaatgaagtgagcgcatcaagctcagccaag	91																					
DB	29	ATGGAAGTGTGTAAGTCTTCACATGAAAGTGCGCATGAGATGCAAGCT-----	80																					

Accession	Position	Sequence	Length
U00001	1	ATGCGAAGCACTCTTCTGCTTCAAAAAGGTTTACTATGCAAAAGCCATTAACGAA	139
U00001	140	GAAGCCACTACTGAACTCTACACCAACTCTTCCCAAAACGATGATGACGACGACT	199
U00001	200	GATTGCTCAAGTGGACCGAACACTCTTACCAATATCTAATGATGATAAAAGTTGAA	259
U00001	260	AAAGTGTGCGCAGGAAGAAGATGAATGAAGTCCCACTCAAGATTTCTCTAAT	328
U00001	329	GAATCTTCCAAATGATTTCAATCGGTTTCAAGTGTGTCGAAGCTCTACCGCAAA	388
U00001	389	CTGAGAAAGAAATGAGCAGAGATGAAGTACGCTGCTTAAGCGCAATGCGTCTC	448
U00001	449	TTCTAAGGAAAGCTCTTCCCAACAAAGCCTCATTTTGTATCATCTTAAGCTC	493
U00001	494	ATGAGCTTTCTCGGTTCTGATTTGGAAAGGTT-----AACAA	555
U00001	556	GGAGATATTACTCTCCAAAGATGTCGCGCGTCCAGAAGCATATTGATGAT	628
U00001	629	TTTCAAGAAAGATTTTACCAATCTTCTGACGAGCCGACGAGAAATTTGTTCCGCGCGG	655
U00001	656	GTTATGCTTAACTTAAATGGAAGAAAGGTGAAGTCAATTCGCGCAAGAGAGTGC	715
U00001	716	TATGCACTTGGAGCTTTAGCCAGGCGCTCAATGATGTTGTTCCAGAGGCGCAATTGAG	775
U00001	776	GAAGAAACAATGATGTTTAAAGTCTCTCAATGACACCATCACACAGAAAGTGA	835
U00001	836	TATTTTGTGGAAGGAGAGATCTTCTCTGAACTGCACTGACGCGCAACAATCTCAC	895
U00001	896	TATGAGCTTATGATATCATGATCATATG--TGACTGTGATCATCATCATGCTTCAAGATGCGG	954
U00001	955	ATACATCTTTTGAAAT--TGCGAGAACGCGGAGGAGAACCTTGTGTTGTTGCTACTTT	1012
U00001	1013	GGAGAAAGCAATTAAGATGAGGTTTCCATAG	1045

DEFINITION	Accession	Version	Keywords	Source	Organism	Reference	Journal	Features
Cucumis sativus EXAF16 mRNA for S-adenosyl-L-methionine:salicylic acid calboxyl methyltransferase-like protein, complete cds.	AB046595	1	AB046595.1	GI:18461099	Cucumis sativus (cultivar:Shimoshirazu-1jibai) shoot apex cDNA, to mRNA	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eustroids 1; Cucurbitales; Cucurbitaceae; Cucumis: 1 (bases 1 to 1308)	Direct Submission	Submitted (26-Nov-2000) Sugihito Ando, University of Tsukuba, Institute of Biological Sciences, Tennoudai 1-1-1, Tsukuba, Ibaraki 305-8572, Japan (E-mail: Sugihito@rs.tsukuba.ac.jp, Tel:81-298-4871(ex.4871), Fax:81-298-4579)
Cucumis sativus (cultivar:Shimoshirazu-1jibai) shoot apex cDNA, to mRNA					Cucumis sativus			Location/Qualifiers
								1..1308
								/organism="Cucumis sativus"
								/cultivar="Shimoshirazu-1jibai"
								/db_xref="taxon:3659"
								/tissue_type="shoot apex"
								32..1144
								/gene="ERAF16"
								32..1144
								/gene="ERAF16"
								/codon_start=1
								/product="S-adenosyl-L-methionine:salicylic acid calboxyl methyltransferase-like protein"
								/protein_id="BAB84353.1"
								/db_xref="GI:18461100"
								/translation="MLHYNGMGNTSYANNSHLOREITSIAKEALTNFYNOHLP
								TSITIDLCSSGQNTLALVYLKIKOEVIRKLRPLEYIFNDLHGDPAVAF
								TSIPRLFELDGIIGIDGPGCFPNVGSFVLRPEKSVPHSSSLHMLSRVPG
								IEKNKNITIGSTSPKSGVAGYRKQKOFMSFLKCAEELWGGSSVLTIGRSD
								PSKSGGYITWELIGLANTVWAGVIEEKNANSPFVITVYIPKREAVYVESFTIL
								NOKASITLNTWYHKREESSPLINSEKLNADADYDFACIOSVSEPLLIRHGDAH
								MDLELFRHNIVACGMKHKRMICMLTISLTKKS"
BASE COUNT	424 a	225 c	257 g	402 t				
ORIGIN								
Query Match	12.28;	Score 157.8;	DB 8;	Length 1308;				
Best Local Similarity	52.28;	Pred. No. 4.5e-30;						
Matches 481; Conservative	0;	Mismatches 392;	Indels 48;	Gaps 4;				
OY 177	acatcaacaagtgatgattaaagttgaggattgagatgcgtttctggacccaacacactt	236						
Db 156	ACATTCCAACTCATCATATGACAGATTTGGGTTGTTCTTCAGCAAAAACACTCTA	215						
OY 237	taacagtcgaggagacatgctgcgaagatgtgcaagaagtcgcgagagaagaatgaat	296						
Db 216	TGCTTTCTTTACTACTGATCAACAAGTCGAAGAAATTCGCC---AAAAGCTTCACCAA	272						
OY 297	tagaagctccaccacttcagattttctgaaatgactttcccaaatgatttcagtcg	356						
Db 273	GGCTGCTTTGGAGATCAACAATTTCTTGAAGATCTTCAGGAATAGCTTCAATCTG	332						
OY 357	tttccaagtcctgcgaagcttaccgcaaacctcgagaagaagaatgacgacgaatag	416						
Db 333	TCTTTACATCATACCAAGTTTTTGAAGATTTGGGACCCAAAATGAGGTGATTTTG	392						
OY 417	gatacgctctaagaagcgaatgcctgcgtcttcttcagcgaacacttcccgagagag	476						
Db 393	GTCCATGTTTCTTCACGAGAGTCCCGGTTCTTCTTACGCCACAGCTTTTCCACACAA	452						
OY 477	ccatgcatattttgcaactctgttacaagtgtcatatggttactcaagttcccaagcgt	536						
Db 453	GTGTTCATTTTTCATATTCATACAGTAGAGCTCCATCATGGCTTTCTTGAGGTTCC-----	504						
OY 537	tgtgtgattgaatggaggatctgtgtgcaacaaggagatattacttctccaaagatgtc	596						


```

Db 505 -----TGTGGGATAGAGAAATACAAAGGAACATTTACAGTAGTACAGTCAAGTC 554
QY gtccgccccccaagaagcatttgatcatcaatttcaagaatttaccacatttca 656
Db 555 CAAGAAGTGTGGGTAGACCTATTATTAAGCAATTCGAAGGATTTTTCATGTTTGA 614
QY ggaatttaccgaagaattgttttcaagtcgcaatgtctcttaccgcaattgttaag 716
Db 615 AGTGTGAGCCAGCAATGTTGTAATGCTGAGAGTATGTTTACATTTGTTGAGAGAA 674
QY 717 tagatg-----aatcgaagcaaccgaacccttagaacttacttgaactgga 764
Db 675 CAAGTAAAGATCCATCAAAATCTGGAGGTACTATATTTGGAGCTTTTGGTTTACCTC 734
QY 765 taaagcaatttgatgttgaaggaacttcggaagaagaanaattgataatttcaatattc 824
Db 735 TCATATCCATGTTGGCTGAGGGAATTTGTGAGACGAAAAAGCGGATTCATTCATTTTC 794
QY 825 catcttaccacccccaagaagaagaatgaagtcagatgttgaagaagaagttcttgcg 884
Db 795 CATACTATATTCATCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
QY 885 aaatttatacttgagaagactttaaagccattatgatgtctcttcttcttgaatg 944
Db 855 TCCTTATCAATTTGAAGCTTCAAGTATTAATTAACCATCTGTCATTAATAAGAGAGAG 914
QY 945 attaccagtaaga-----tccatgaacaatttaagcagaagatgtg 989
Db 915 AAGATCAAGTACTCCACTAATTAATTAATTCCTGCGCCATCTCTCATCTATGATTTG 974
QY 990 calcataattgatacagtttaagaaccatctctgcgaagcatttggagaagcattta 1049
Db 975 CCAATGCTATCAATCTGCTGAGACCACTTTGATGCTGATTTGAGAGAGAGAGAGAG 1034
QY 1050 tgcctgaacttattccacagc 1070
Db 1035 TGGATGATTAATTAATTCAGAC 1055

RESULT 8
AB031280 1438 bp mRNA linear PLN 31-AUG-2000
LOCUS AB031280
DEFINITION Camellia sinensis TCSI mRNA for caffeine synthase, complete cds.
ACCESSION AB031280
VERSION AB031280.1 GI:9967142
KEYWORDS caffeine synthase.
SOURCE Camellia sinensis CDNA to mRNA.
ORGANISM Camellia sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Ericales; Theaceae; Camellia;
1 (sites)
REFERENCE
AUTHORS Kato,M., Mizuno,K., Crozier,A., Fujimura,T. and Ashihara,H.
TITLE Caffeine synthase gene from tea leaves
JOURNAL Nature 406 (6793), 936-957 (2000)
MEDLINE 20437335
REFERENCE
AUTHORS Kato,M., Mizuno,K., Ashihara,H. and Fujimura,T.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1999) Kouchi Mizuno, University of Tsukuba,
Institute of Agricultural and Forest Engineering, 1-1-1 Tennoudai,
Tsukuba, Ibaraki 305-8572, Japan
(E-mail:koumoe@akura.cc.tsukuba.ac.jp, Tel:81-298-53-4656,
Fax:81-298-55-2203)
FEATURES
source 1. 1438
Location/Qualifiers
/organism="Camellia sinensis"
/db_xref="taxon:4442"
90..1199
/gene="TCSI"
90..1199
CDS
/gene="TCSI"

```

```

/function="N-methyltransferase"
/codon_start=1
/product="caffeine synthase"
/protein_id="BAB12278.1"
/db_xref="GI:9967143"
/translation="METATAGCAATVETLPMNRGSESSVAONSSFTQVAMAPALEN
AVETLSPRDFHIALNADLCAATVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEV
DLRGNPNTLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEV
LTPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEV
CMVLTIRGQCSDEPDMSCSTWELMALIAELVSGLLIDEDKLTETLSPRSEVETLSPRSEV
KQIVRDSFTLIDISFDLDSVEMENKVRGKFTKVRATETLSPRSEVETLSPRSEVETLSPRSEV
MDVLYKFTLHIVSDLEAKFTLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEVETLSPRSEV
BASE COUNT 453 a 248 c 326 g 411 t
ORIGIN
Query Match 10.8%; Score 140.4; DB 8; Length 1438;
Best Local Similarity 51.9%; Pred. No. 1.4e-25;
Matches 462; Conservative 0; Mismatches 391; Indels 37; Gaps 5;

QY 34 ggaagtcgaagaatctctgcatatgaatgaagtgagagagatcaagtcagcaaga 93
Db 110 GAAGGTAAAGCAAGTGTGTTGATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 169
QY 94 tgcattccacacatctgctcttgc-----caagtgaaaccttctcttgaagatgata 151
Db 170 CTCTTCTTACAGCAACAAAGTGCCTCAATGCGCACAGCCAGCCTAGAAAATGCAAGTTCA 229
QY 152 gaattgttgaggcgaacttgcacacatcaagaatgacatgaatgttgcgaattggga 211
Db 230 AAGTCTCTTCTCCAGAAATTTCC-----ACCTTCAAGCTTTTACGAGCGGACTTGGGT 284
QY 212 tgcgcttcggaacacaacacactttaaagtcggaacattgtgcaaatgatacaaa 271
Db 285 TGTGAGGCGGTCCAAACACATTCGACATGATTTCTACGATCAAGAAAGATGAGAGAAAG 344
QY 272 gtggccagagaagaagaatgaatgaagtcgacacacacacacacacacacacacacac 331
Db 345 AATGTCAGAGAAATGA-----ATTGCCAAACACTGGAACCTGAGTTTACTTGAAATGAT 398
QY 332 ctttccaaatgattcaatcggttccaagttgtcgcgaagcttaccgcaaac 391
Db 399 CTTTGTGAATGATTTCAATACCTCTTCAAGGCTGTGCTGTGAG----- 446
QY 392 gagaagaagaatgagcagaagatagatgttgcctaaataagcgaatgctgtcttcc 451
Db 447 GTTATGTAAACAATGTGAGAAAGTCCGTGTTATGTGATGAGGAGTACCGGCGCTTTC 506
QY 452 tgcgagacacttcccgagagagatccatgacatttttgacatttgaagtgat 511
Db 507 CATGCCCGCTTTTCTCTGTACAGCTTACATTTAGTTCACTCTCTACAGTGTCTAT 566
QY 512 tggattatccaggttccagcgggttggatgattggaggattgagcaacaagaag 571
Db 567 TGGCTTACTCAGGCAACCAAGAGACTCACAAACAGAGAGAGGCTTGGCATTTAAACAAGGG 626
QY 572 agtattacttccaaagatgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 631
Db 627 AAGATTTACATATCAAAAGCAACACCCCTCTGTTGTAAGAGAGAGCTTATCTACATTT 686
QY 632 acgaagaatttaccacatttcaagaatttcaagaagatgttgaagtgatgagcga 691
Db 687 CATGAAGATTTCAAAAGTGTTCATATGCTAGATCCCAAGAGGAGGAGTCCAAATGTTGT 746
QY 692 agtccttaccatcttgaatgagatgagatgagatgagatgagatgagatgagatgagat 739
Db 747 ATGAGTGTGATCTGCTGTGAGCAATGTTCTGATCTGATGATGATGATGATGATGATGAT 806
QY 740 ccccaagacttacttgaatgagatgagatgagatgagatgagatgagatgagatgagat 799
Db 807 ACTTGGAGACTTATGATGAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 866
QY 800 gaaaatgagatgattcaatattccatttaccaccttcagcagaagaagaatgagc 859

```

```

Db 867 GATAAATTGACACCTTCAATATACCCAGTATTTTGCATCTGCTGAGAGAACTCAATAGAT 926
OY 860 atagttgaggaagagttcttcggaatttatcttcggaacttttaa 909
Db 927 ATATGTGAGAGGAGGAGGATTCATTCATATGATCATATGAGGGTTGA 976

RESULT 9
AX138776 AX138776 1427 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 2 from Patent EP1055727.
DEFINITION AX138776
ACCESSION AX138776
VERSION AX138776.1 GI:14274532
KEYWORDS
SOURCE
ORGANISM
Camellia sinensis.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Ericales; Theaceae; Camellia.
1 (bases 1 to 1427)
REFERENCE
Mizuno, M., Ashihara, H., Mizuno, K. and Fujimura, T.
Cloning of an n-methyltransferase involved in caffeine biosynthesis
Patent: EP 1055727-A 2 28-NOV-2000;
JOURNAL MITSUI CHEMICALS, INC. (JP)

FEATURES
source
1. 1427
/organism="Camellia sinensis"
/db_xref="taxon:4442"
BASE COUNT 444 a 251 c 330 g 402 t
ORIGIN

Query Match 10.7%; Score 138.8; DB 6; Length 1427;
Best Local Similarity 51.8%; Pred. No. 3.7e-25;
Matches 461; Conservative 0; Mismatches 392; Indels 37; Gaps 5;

OY 34 ggaagtcacaagaagtcctgcataatgaatgaagtgagagcgatatacagccaagaa 93
Db 80 GAAGTGAAGCAAGAGTGTTCATGAAACAGGGGGGAGGAAGAAAGTATGCAACAAA 139
OY 94 tgatcctcaaatcctgcctctgc--caagtgaaaccttcttgaacaatgcatacga 151
Db 140 CTCTCTTTCAGCAACAAAGTGCGCTCATATGCGACAGCGGCTTGAATAATGCTGA 199
OY 152 gaattgttggcggaacttgcacacatcaacaagtgcattaaagttggagttgga 211
Db 200 AACTCTCTTCACAGATTTTC-----ACCTTCAAGCTTTTACGACAGGACTTGCT 254
OY 212 tgcgctcttgagcaacaacacatttaacagtgcggaagcattgcaaatgacaa 271
Db 255 TGTGCAAGCGGTCACAAACATGCGAGTATTTTACGATCAAGAAATGATGGAAGA 314
OY 272 gtggccaggaagaagaatgaatgaatgaacgctccacacatttaagaattttcgaatgat 331
Db 315 AAATGCAAGGAATTTGA-----ATTGCCAAACACTGGAACCTTCAAGCTTTACTGAAATGAT 368
OY 332 ctcttcaaaaatgatcattcattcggttttcaagttgtgcgaagcttcaaccgcaaatc 391
Db 369 CTTTGTGAAATGATTTCAATCCCTCTTCAAAAGCGCTGCGCTGAG-----416
OY 392 gagaagaagaatgaggaagaatgagatcgctgcttaataagagcgaatgcgtcttc 451
Db 417 GTTATGTGTAACAAATGTAGAGAAATCCGCTGTATGTATGTAGAGATACCGGCGCTTTC 476
OY 452 taaggagactctcccggaagagtcacatgatttttgaactctgttaagatgtat 511
Db 477 CATGGCGGCTTTTCTCTGTATACAGCTTACATTTAATTCATCTCTTACAGGTTTAT 536
OY 512 tggttatcctaggttccagcggttggtagctgaattggaggtctgggcaacaagag 571
Db 537 TGGCTTACTACGACCAAAAGGACTCACAGACAGAGAAGGCTTGGCATTTAAACAAGGG 596

```

```

OY 572 actattactcttccaaagagatgctgcgcccgtccaggaagacatttgatcaatt 631
Db 597 AGATTTTCATATCAAAAGCAAGCCCTCTGTGTAGAGAGAGCTACTATCTCAATT 656
OY 632 acgaagaatttaccacatttcaagatcattcgaagaagtggtttcaagtgccga 691
Db 657 CATGAAGATTTCACAAATGTTTCTCATATGCTAGATCCCAAGAGTGTTCAATAGTGT 716
OY 692 atgcctcttaccgtcatttgtaagtagat-----gaatcgagaaacgaat 739
Db 717 ATGGGTGTAATCTTCGTGTAGGCAATTTCTGATCCCTTCACATGCAAGAGCTGCTTT 776
OY 740 cccctagacttactgcacatgcgaataaagacgttgattgttgaggaacttcgagagaa 799
Db 777 ACTTGGAACCTATTAGCTATGCTATGCGCATTTGCTGTAATGTTTCACAGGATTTGATGAA 836
OY 800 gaaaattggaatgttccatattccattcttcaacattcagcagaagaagaatgctc 859
Db 837 GATTAATTTAGACACCTTCATATATACCCAGCTATTTTGCATCATCTTGAGGAAGAAAGAT 896
OY 860 atagttgaggaagaggttcttcggaatttatcttcggaacttttaa 909
Db 897 ATATGTGAGAGGAGGAGGATTCATTCATATGATCATATGAGGGTTGA 946

RESULT 10
AX138777 AX138777 1427 bp mRNA linear PAT 30-MAY-2001
LOCUS Sequence 3 from Patent EP1055727.
DEFINITION AX138777
ACCESSION AX138777
VERSION AX138777.1 GI:14274533
KEYWORDS
SOURCE
ORGANISM
Camellia sinensis.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Ericales; Theaceae; Camellia.
1 (bases 1 to 1427)
REFERENCE
Mizuno, M., Ashihara, H., Mizuno, K. and Fujimura, T.
Cloning of an n-methyltransferase involved in caffeine biosynthesis
Patent: EP 1055727-A 3 29-NOV-2000;
JOURNAL MITSUI CHEMICALS, INC. (JP)

FEATURES
source
1. 1427
/organism="Camellia sinensis"
/db_xref="taxon:4442"
BASE COUNT 444 a 251 c 330 g 402 t
ORIGIN

Query Match 10.7%; Score 138.8; DB 6; Length 1427;
Best Local Similarity 51.8%; Pred. No. 3.7e-25;
Matches 461; Conservative 0; Mismatches 392; Indels 37; Gaps 5;

OY 34 ggaagtcacaagaagtcctgcataatgaatgaagtgagagcgatatacagccaagaa 93
Db 80 GAAGTGAAGCAAGAGTGTTCATGAAACAGGGGGGAGGAAGAAATGATTTGCAACAAA 139
OY 94 tgatcctcaaatcctgcctctgc--caagtgaaaccttcttgaacaatgcatacga 151
Db 140 CTCTCTTTCAGCAACAAAGTGCGCTCATATGCGACAGCGGCTTGAATAATGCTGA 199
OY 152 gaattgttggcggaacttgcacacatcaacaagtgcattaaagttggagttgga 211
Db 200 AACTCTCTTCACAGATTTTC-----ACCTTCAAGCTTTTACGACAGGACTTGCT 254
OY 212 tgcgctcttgagcaacaacacatttaacagtgcggaagcattgcaaatgacaa 271
Db 255 TGTGCAAGCGGTCACAAACATGCGAGTATTTTACGATCAAGAAATGATGGAAGA 314
OY 272 gtggccaggaagaagaatgaatgaatgaacgctccacacatttaagaattttcgaatgat 331
Db 315 AAATGCAAGGAATTTGA-----ATTGCCAAACACTGGAACCTTCAAGCTTTACTGAAATGAT 368

```


	RESULT	12	
DB	BD004679	1427 bp	RNA linear
	LOCUS	Gene encoding caffeine synthetic system-related enzyme and utilization thereof.	PAT 31-JAN-2002
	ACCESSION:	BD004679	
	VERSION:	JP 2001037490-A/2.	
	KEYWORDS:	Camellia sinensis.	
	SOURCE:	Camellia sinensis	
	ORGANISM:	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteriales; Ericales; Theaceae; Camellia.	
	REFERENCE:	Mizuno,M., Ashihara,A., Mizuno,K. and Fujimura,T. Gene encoding caffeine synthetic system-related enzyme and utilization thereof Patent: JP 2001037490-A 2 13-FEB-2001;	
	AUTHORS		
	TITLE		
JOURNAL			
COMMENT			
	OS	Camellia sinensis	
	PN	JP 2001037490-A/2	
	PD	13-FEB-2001	
	PF	23-MAY-2000 JP 2000151718	
	PI		
	PC	MISAKO MIZUNO,AKIRA ASHIHARA,KOICHI MIZUNO,TATSUTO FUJIMURA C12N5/09,A01HS/00,C12M/15,C12N1/19,C12N1/21,C12N5/10, PC C12N5/10,C12N9/10, PC C12P17/18,C12P23/00//C12N1/21,C12R1:19),C12N5/10,C12R1:91), PC C12N5/00, CC C12N5/00,C12N5/00),(C12N5/00,C12R1:91)	
	FC		
	CC		
	FT	Key	Location/Qualifiers
	FT	source	1..1427
FEATURES			
source		Location/Qualifiers	/organism='Camellia sinensis'
			1..1427
			/db_xref='taxon:4442'
BASE COUNT	444 a	251 c	330 g 402 t
ORIGIN			
	Query Match	10.7%; Score 138.8; DB 6; Length 1427;	
	Best Local Similarity	51.8%; Pred. No. 3.7e-25;	
	Matches 461; Conservative	0; Mismatches 392; Indels 37; Gaps 5.	
OY	34	ggagcccaagaagctcgtgcatatgaatgaagtgaagtgacatacaagctaacgcacaagaa	93
DB	80	GAAGCGAACAAGGATTGTTCATGAAACGGGGGAGAAGAAAGTATGACACAAA	139
OY	94	tgatctctaacttcggctctgac--caagtgtaaacotttcttgaacaatgatacga	151
DB	140	CCTTTCTTTCACGCCAACAGTGCGCTCATATGCACACGACGCGTAGAAAAATGCAATTGA	199
OY	152	gaatttgtcgaggccaactggccaacatcacatacaagtgacatttaagttagtgattgga	211
DB	200	AACCTCTTCTCCAGAGACTTTC-----ACCTCAAGCTCTTAACCCACAGGACTGTGGT	254
OY	212	tgcgtcttgagcaaaaacacctttaaacagtgagggagcaatttgcaagaatttaaaaa	271
DB	255	TGTGCACCGGGTCCAAACACTTTCGACGTGATTTCTACGATCAAGAAATGATGGAAG	314
OY	272	gttgccaggaagaagaagaatgaatagaagctcccaccattcaagattttctgaagat	331
DB	315	AAATGACGGGAATTAA-----ATGACCACAAACACGTGGAACCTCAGGTACTTGATAT	368
OY	332	ccttccccaaatgatctcaattcggitttcaagtgtgtgcacaagcttctacgccaaact	391
DB	369	CTTTTGGAAAGATTTCATAATACCTCTTCAAAGGCGTGTCTGAG-----	416

Oy		392	gagaagaanaatggacgaagataggatcgcgtccaaataaagcgcaatgccgtgctttc	451
Dd		417	GTATTGGTACCAATGTGAGGAAGTCCGGTTAATGTATGTATGGAGTCACCGGGGTCTTTC	476
Oy		452	taccgacaccttcccgcggagtcgatgatatttttgactctgtttaagtgtcat	511
Dd		477	CATGCCCGCTTTTCTCCTGTAAACAGCTTAATTACTTATTCATTCTCTTAAGTGTCAT	536
Oy		512	tgttatcttagtctcccgaggttgttgatgaattgsgatttgyscacaaaagg	571
Dd		537	TGGCTACTCAGGCACCMAAGACTCACAGCAGAGAAGCGTTGGCATTAACAAGGG	596
Oy		572	agattactcttccaaggagtgctgcccccgtccagaaggacaatttgatcaatt	631
Dd		597	AAGATTTCATATTAAGAACACAGCCTCCTGTGTGAAGAGACCTACTTATCTCATTT	656
Oy		632	acgaagaatttacacacattctaaggatcattcagaagaagittgttttcacgtgcga	691
Dd		657	CATGAAGATTTCACAAAGTTTCTCAATGCTAGATCCAMAGAGGTGGTTCCAATAGTGTG	716
Oy		692	atgctcttaccctgcatcttgtaagtagat-----gaattcgagaacgaat	739
Dd		717	ATGGCTGTGAATACTCTGTGTAGCGCAAGTTCTGATCCTTCACACATCAAGTGCCTTT	776
Oy		740	ccccctgaccttactgtgacagygaaataaagacttgatttgtagggactcttgagaa	799
Dd		777	ACTGTGAACATTAATAGCTAATGGCAATGCTGTAATATGGTTCCAAAGGAGTATGATGA	836
Oy		800	gaaaaattggatgtttcattcattcttcattcttccacttcacgttcagcagaagaagtaa	859
Dd		837	GATTAATTAGACCTTCATATFACCGATATTTTCGACACTTGAGACCTTGAGAAAGAT	896
Oy		860	atagtgtgagagaaagtcttgsgaaattttatatctgsgaatctttaa	909
Dd		897	ATAGGAGAGGAGCGAGCATTCATCACATTTGATCATTAAGAGGGGTTTTGA	946
RESULT	13			
LOCUS	AF198492		1363 bp	mRNA linear PLN 10-AUG-2000
DEFINITION	Antirrhinum majus SAM:benzoic acid carboxyl methyltransferase (BANT) mRNA, complete cds.			
ACCESSION	AF198492			
VERSION	AF198492			
KEYWORDS	AF198492.1 GI:9789276			
SOURCE				
ORGANISM	snapdragon.			
REFERENCE	Antirrhinum majus Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.			
AUTHORS	1 (bases 1 to 1363) Dudareva,N., Murfitt,L.M., Mann,C.J., Gorenstein,N., Kolosova,N., Kish,C.M., Bonham,C. and Wood,K.			
TITLE	Developmental regulation of methyl benzoate biosynthesis and emission in snapdragon flowers			
JOURNAL	Plant Cell 12 (6), 949-961 (2000)			
MEDLINE	20012758			
REFERENCE	2 (bases 1 to 1363)			
AUTHORS	Dudareva,N., Murfitt,L.M. and Mann,C.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-OCT-1999) Horticulture and Landscape Architecture, Purdue University, 1165 Horticulture Bldg., West Lafayette, IN 47907-1165, USA			
FEATURES				
source	location/Oualifiers			
	1..1363			
gene	/organism="Antirrhinum majus"			
	/db_xref="taxon:4151"			
CDS	1..1363			
	/gene="BANT"			
	19..1113			
	/codon_start=1			
	/product="SAM:benzoic acid carboxyl methyltransferase"			

```

/protein_id="AAF98284.1"
/db_xref="gi:9789277"
/translation="MKYAKKILKININAGDESVYANNGLQKVMKSLHYLDELKLD
IGDHYGPRCKKMDKSGSPMALIVSGIINTEDYIKENINELPEFEVFLNDL
PMDNFNLEFKLHSHENGKCYVGLPSGFGRLPKSKLHFVAYSSJHMSOYEGLE
DNDRORIVATPEPEYKAYAKOYERSTFKLKEGMYLTVGNGSVSDP
SSADDLAFTLTLKIVDMVAEYKMKDDLYSPKPIYSPCEKREVAALISEGFTLD
RLFEVPCVMDSDYTDODDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
LRGKAKKRVKVEHLSEVNSYFSTVLSRR"
BASE COUNT      417 a      239 c      300 g      407 t
ORIGIN
Query Match      10.4%; Score 134.8; DB 8; Length: 1363;
Best Local Similarity 49.6%; Pred. No. 4e-24;
Matches 546; Conservative 0; Mismatches 482; Indels 72; Gaps-
5;
QY 59 aatgaagtgaaagcgatatacaagctacgcaagaatgcatctcactgctctgccc 118
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 52 ATTGCAGAGATGTTGAAACTGATCTACGCCAACAAATTCGCTTCAAAAAGTTATGATG 111
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 119 aaggtgaacaccttcctgacaacatgcatagagaattgttcgagcgcaactgcccac 178
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 112 TCAAAATCATTCATGTTTATGACGAACCCCTTAAGATATTCGGTGCATGTTGGC 171
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 atcaacaagtcgcatgaagttcgagttcgagtcgctctgaccacaacacttta 238
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 172 TTCCTCAAAATCTTCAGATGATGATGATGGTGTTCATCAGGCGCTACGCCCTTTTG 231
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 239 acaagtcgagcactgtgcgaagattgcaagaagtcgcagagagaagaatgata 298
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 232 GTCATGTCGGCATTATTAATTAATGAGATTTT---GTACAGAGAGAGATATATAT 288
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 299 gaagtcgccacactgagatttttcgaatgactttcccaaatgattcaattcggt 358
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 289 GAATTAAGTGAATTTGAGGTTTTCGAACGATCTTCAGACAAACGATCAACACCTC 348
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 359 tcaagttgcgcgaacactctccgcaaacctcgagaagaatgagcgaataga 418
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 349 TTCAATATGTTATC-----ACATGAGATGGA 375
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 419 tcatgctaataaagcgaatgctgctctctctacggcagaactctcccgagagctcc 478
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 376 AACTCTCTTGTATATGTTGCTGTGATCTTTCTACGGGAGACATATTGCCAAAAAGAGC 435
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 479 atgcatctttgcaactctgttacaagtgatcattgattcagagttcccgagcttg 538
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 436 CTACACTTTCCTATTCCTTCACAGATTCACGTGCTCTCAGAGTTCTCGAAGGCGTG 495
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 539 gtagtggaattgggagattgttgcaaaaaggagattattactctccaaaagattgct 598
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 496 GAGGATAT-----MACAGACAAACATTTCATGAGGACGAGAAATCCT 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 599 tgcgcgcgcgaagaagcatatttgatcaattacgaagaatttaccacatttcaag 658
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 541 CCGGAAGTGTCAAAAGCATACGCAAAAGCAATTCGAAAGACATTCACCAATTTCTAAAG 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 659 attcatctgaagaagattgttttcaagtcggaagtccttaccctgacttgaagta 718
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 601 TTGCGAGCGGAGAAATTTGACAGAGTGAGACCATGCTCTTGACATTTAAACGCGAGAACT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 719 gatga-----ttcgagcaaccgaatccccaagacttaacttgacatgagcaata 766
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 661 GTTGAAGATCCCTCGAGCAAAAGATGACTTACATTTTACATTTCTGCAAAAACGCTA 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 767 aagaagtgattgttgaggaactctcgagaggaagaaatagttgagtagttccaatctca 826
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 721 GTTGAATATGTCGTCGAGGCGCTTTCACAGATGAGCAATTTGTACTCGTTAACATTTCT 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 827 ttcttaagccttcagcgaagaagtaagtcataagttgaggaaggaagttcttgcaaa 886
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 781 ATTACTACACATGTACCGCGGAGTAGAGGACACGCAATTTCTGATGAGAGGCTCTTTTACG 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 887 atttatactggaagactttaagcccatatgactgctgctctctattgatgat 946
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 841 TTGGACAGGCTAGAGAGCTTTCTGCTTTGTTGGAGTCAAGTACATACAGAGAGAT 900
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 947 -----taccagtaagatcccatgaaacaatlaagcagaatgctgcatca 997
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 901 GATCAGCAAGACCCATCATCTTTGGCAAAAGAGTGAATAATTTGTGCGAGATTTG 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 998 attagatcagtttaagaaaccatctcctcgcaagtcattttggaagaagcattatgcctgac 1057
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 961 GTACGGCTATTACGGAACCAATGCTGTAGCATTTTGGAGCAGCATATTATGATCTT 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1058 ttatccagagcttgcgaagcagatgacgaagaagttctccacttgaggaagagctgctat 1117
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1021 CTATTGTGAAGATGCCAAAGAAATTAAGTGCAGCATCTATCTGTGAGAACTCGCATAT 1080
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1118 aatactctatcattctct 1137
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1081 TTCAGCATACTACTTCTCT 1100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 14
AF179222 1306 bp mRNA linear PLN 06-JUN-2000
LOCUS
DEFINITION
Brassica rapa subsp. pekinensis floral nectary-specific protein
(Ntr1) mRNA, complete cds.
ACCESSION
AF179222
VERSION
AF179222.1 GI:6651394
KEYWORDS
SOURCE
ORGANISM
Brassica rapa subsp. pekinensis.
Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 1306)
Song,J.T., Seo,H.S., Song,S.I., Lee,J.S. and Choi,Y.D.
Ntr1 encodes a floral nectary-specific gene in Brassica campestris
L. ssp. pekinensis.
Plant Mol. Biol. 42 (4), 647-655 (2000)
JOURNAL
MOLBIOL
PUBMED
10809010
2 (bases 1 to 1306)
Song,J.T., Seo,H.S., Song,S.I., Lee,J.S. and Choi,Y.D.
Direct Submission
Submitted (20-AUG-1999) Agricultural Chemistry, Seoul National
University, Seodun-dong, 103, Suwon, Kyunggi-do 441-744, Korea
FEATURES
Source
1. 1306
/organism="Brassica rapa subsp. pekinensis"
/sub_species="pekinensis"
/db_xref="taxon:51351"
1. 1306
/gene="Ntr1"
4. 1182
/gene="Ntr1"
/feature="Ntr1"
/codon_start=1
/product="floral nectary-specific protein"
/db_xref="gi:6651395"
/protein_id="AAF22289.1"
/translation="MEWRILHNKNGETSYAKNSIVQSNITSLGRRVDEALKLIM
IRNSLSFGIADLGCSGPNLSLSINIVETIONLCHDLRVPPELSLNDLPND
FNITFASLPEFEDRYKRRDNVYSLGFEHSGGCPFVSAGSYGRLFPRLSHFV
SSSLHMLSOVPCGKVNKDGVTADLDNRGKIVLSLTSRSLAKYVALQFOTDFSV
FLRSSEELVPGGRVLSFLGRSSPDPTTEESYOMELLADLAKKEGIIIEENID
AFAPAPYASPEELKMALEKGSFSDIRLEISVWMEGGSISDSYDIVRKEALAS
GRVAKTIRAVVEPLEPTFGOKYMDLEFERAKIVGEVYVSSPRYTIIVLSLRNG"
BASE COUNT      367 a      267 c      308 g      364 t
ORIGIN
Query Match      8.4%; Score 108.4; DB 8; Length 1306;

```



```
OY 737 aatcccttagacttacttgac-----atggcaataaagacttgattgttgaag 785
DB 761 CTGAATGCTGTCTCATTTGGCAACTCTAGCCATGCGCTCTCAATCAATGCTTTCGAGG 820
OY 786 gacttccttgagagagaaataatggaatggttcaatattccattcttaacacctcagcag 845
DB 821 GACTTAATGAGAGAGAGAGAGATGATTAAGTTCAACATTCCTCAGATACACCATCTCCAA 880
OY 846 aagaagtaagtgatgcttgagagagaggtctctgcgaatttatatcttgagactt 905
DB 881 CAGAACTAGAACAGAGATCCCTAAAGAGGGCTTTTGTGATGACCATATAGAGGCTT 940
OY 906 ttaaggcccatatgatgctgctctctcattgatgatattaccagtaagaatcccatg 965
DB 941 CAGAAATATACTGAGAGT-----AGCTGCACCTAAAGATGCTGATGCTGGCTGTG 994
OY 966 aacaattaaagcagagatgctgcatcatttaattagatcaggtttaaaccatctcgcg 1025
DB 995 AGGAA---GAAAGTTACAGAGTGGCTCGGTCATGAGAGAGTGGCCGAGCCATGCTGC 1051
OY 1026 caagtcatttggagagagctattatgctgacttattcacaagcttgcgaagcagcag 1085
DB 1052 TCGACATTTTGTGTAAGCCATCTGTAAGATGTGTTCCATATGGATTAACCTGCTCTTA 1111
OY 1086 caaagttctccacatgagcagaagctgctataataatctatcatctctctgcgaaga 1145
DB 1112 TCGAAAGAAATGCTTAAGAGAGAACCAANTCATCAAGCTCATGTCTCTCATTCGAA 1171
OY 1146 agccagagaagt 1157
DB 1172 AATCAAGATTAAAT 1183
```

Search completed: July 27, 2002, 04:39:42
Job time: 4583 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 27, 2002, 03:26:19 ; Search time 54.13 Seconds
(Without alignments)
5890.122 Million cell updates/sec

Title: US-09-971-020-2

Perfect score: 1298
Sequence: 1 agcagcgcgaatcgtatgt.....ataaaaaaaaaaaaaa 1298

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/5E.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/5F.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	3.2	1299	US-08-785-076-1	Sequence 1, Appl
2	39.8	3.1	3839	US-09-056-105-14	Sequence 14, Appl
3	38.2	2.9	991	US-08-924-747-25	Sequence 25, Appl
4	38.2	2.9	991	US-09-247-3738-25	Sequence 25, Appl
5	37.2	2.9	991	US-09-296-715-25	Sequence 25, Appl
6	36.6	2.8	120	US-07-867-106-2	Sequence 2, Appl
7	36.6	2.8	120	US-08-153-0518-28	Sequence 44, Appl
8	36.6	2.8	120	US-08-060-952C-44	Sequence 28, Appl
9	36.6	2.8	120	US-08-151-477A-28	Sequence 58, Appl
10	36.6	2.8	120	US-08-819-867-58	Sequence 88, Appl
11	35.8	2.8	1316	US-09-517-039-88	Sequence 8, Appl
12	35.8	2.7	3205	US-09-086-768A-3	Sequence 16, Appl
13	35.2	2.7	198	US-08-330-108-16	Sequence 16, Appl
14	35	2.7	198	PCT-US92-10087-16	Sequence 16, Appl
15	35	2.7	1915	US-09-120-365-2	Sequence 2, Appl
16	34.8	2.7	1915	US-09-515-039-2	Sequence 2, Appl
17	34.8	2.7	2238	US-08-617-860B-28	Sequence 28, Appl
18	34.6	2.7	2238	US-08-553-619B-8	Sequence 10, Appl
19	34.6	2.7	1174	US-07-869-933-10	Sequence 10, Appl
20	34.4	2.7	1174	US-09-103-663-10	Sequence 5, Appl
21	34.4	2.7	1446	US-08-596-024-5	Sequence 5, Appl
22	34.4	2.7	1446	US-09-020-818-5	Sequence 5, Appl
23	34.4	2.7	1446	US-08-907-740-5	Sequence 5, Appl
24	34.4	2.7	2334	US-08-062-632-4	Sequence 4, Appl
25	34.2	2.6	1089	US-09-004-731-15	Sequence 15, Appl
26	34.2	2.6	1089	US-09-004-731-17	Sequence 17, Appl
27	34.2	2.6	1089	US-09-004-731-17	Sequence 17, Appl

28 34.2 2.6 1089 4 US-08-749-699-15
29 34.2 2.6 1089 4 US-08-749-699-17
30 34.2 2.6 2606 1 US-08-408-318-1
31 34.2 2.6 2606 1 US-08-839-164-1
32 33.8 2.6 1454 2 US-08-522-421-2
33 33.8 2.6 1454 2 US-08-713-000-7
34 33.8 2.6 1454 2 US-08-975-316-7
35 33.8 2.6 1454 2 US-08-975-316-7
36 33.8 2.6 1454 2 US-08-975-316-7
37 33.8 2.6 3501 1 US-07-977-451-3
38 33.8 2.6 3501 1 US-08-252-517-3
39 33.8 2.6 3501 1 US-07-906-397A-3
40 33.8 2.6 3501 1 US-08-601-891-3
41 33.8 2.6 3501 2 US-09-021-324-3
42 33.8 2.6 3501 5 PCT-US92-05401-3
43 33.8 2.6 3501 5 PCT-US92-09893-3
44 33.6 2.6 575 1 US-08-554-659-11
45 33.6 2.6 1166 5 PCT-US96-12129B-1

ALIGNMENTS

RESULT 1
US-08-785-076-1
Sequence 1, Application US/08785076
Patent No. 5789217
GENERAL INFORMATION:
APPLICANT: Hodgson, John
TITLE OF INVENTION: No. 5789217el tRNA Synthetase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,076
FILING DATE: 17-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9601096.2
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 961845.6
FILING DATE: 27-JUL-1996
APPLICATION NUMBER: 9622617.0
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P13354-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1299 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-785-076-1

APPLICANT: MCGONIGLE, BRIAN
 APPLICANT: O'KEEFE, DANIEL
 TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
 FILE REFERENCE: CL-1108-A
 CURRENT APPLICATION NUMBER: US/09/247,373B
 CURRENT FILING DATE: 1999-02-10
 PRIOR APPLICATION NUMBER: 08/924,747
 PRIOR FILING DATE: 1997-09-05
 NUMBER OF SEQ ID NOS: 56
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 25
 LENGTH: 991
 TYPE: DNA
 ORGANISM: SOYBEAN

US-09-247-373B-25

Query Match 2.9%; Score 38.2; DB 4; Length 991;
Best Local Similarity 55.7%; Pred. No. 0.26; Mismatches 58; Indels 0; Gaps 0;
Matches 73; Conservative 0;

OY 1168 aaagttggttttaagttggttttgcgcgttggtgggtcttcgtggtatgtcgtttgt 1227
DB 839 aaagattgttcgttttaatttaattgaactgtggttgcgttgcatttca 898
OY 1228 attcgttaataaagtgatgcaagaataagatatattagtaacataatttcataaaaaa 1287
DB 899 atttaactaaaaaagtggttcgtttcaaaaaaaataaaaaaaataaaaaa 958

OY 1288 aaaaaaaaaa 1298
DB 959 aaaaaaaaaa 969

RESULT 5

US-09-296-715-25
Sequence 25, Application US/09296715
Patent No. 6171839

GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
TITLE OF INVENTION: ENZYMES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/296,715

FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CL-1108
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 991 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: SOYBEAN
TISSUE TYPE: SOYBEAN
IMMEDIATE SOURCE:
CLONE: SSM, PR0067, G5
US-09-296-715-25

Query Match 2.9%; Score 38.2; DB 4; Length 991;
Best Local Similarity 55.7%; Pred. No. 0.26; Mismatches 58; Indels 0; Gaps 0;
Matches 73; Conservative 0;

OY 1168 aaagttggttttaagttggttttgcgcgttggtgggtcttcgtggtatgtcgtttgt 1227
DB 839 AAAGATTGTTCTGTTTAAATTAATTGAATGACTGTGATTTGGTTGGGTAATGCAATTTTA 898
OY 1228 attcgttaataaagtgatgcaagaataagatatattagtaacataatttcataaaaaa 1287
DB 899 ATTTTAACCTAAAAAAGCTTCACGTTTAAAAAATAAAAAAATAAAAAAATAAAAAA 958

OY 1288 aaaaaaaaaa 1298
DB 959 AAAAAAAAAA 969

RESULT 6

US-07-867-106-2/c
Sequence 2, Application US/07867106
Patent No. 5389526

GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Silme Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625

PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feehey, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3439
TELEFAX: 215-568-3100

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 2.9%; Score 37.2; DB 1; Length 5852;
Best Local Similarity 56.6%; Pred. No. 1;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 1177 tttaagttggttttgcgcgttggtgggtcttcgtggtatgtcgtttgtatcgttaat 1236
DB 1177 tttaagttggttttgcgcgttggtgggtcttcgtggtatgtcgtttgtatcgttaat 1236

```

Db      30 GAAATTACTTCATCATGTAGAGAAATAAACAATTGGCACATTTGTAAAAAAAAAAAAAAAA 85
QY      1292 aaaaaaa 1298
        |||||
Db      90 AAAAAA 96
RESULT 8
JS-08-060-95XC-44

```

```

RESULT 8
US-08-060-952C-44
Sequence 44, Application US/08060952C
Patent No. 5695932
GENERAL INFORMATION:
APPLICANT: Michael D. West
APPLICANT: Jerry W. Shay
APPLICANT: Elizabeth Blackburn
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF CONDITIONS
TITLE OF INVENTION: RELATED TO TELOMERE LENGTH AND/OR
TELOMERE ACTIVITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 Diskette, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 07/882,438
FILING DATE: May 13, 1992
PRIORITY DATE: May 13, 1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/882,438
FILING DATE: May 13, 1992
APPLICATION NUMBER: 08/038,766
FILING DATE: March 24, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Waiburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 202/045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEO ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 120
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-060-952C-44
Query Match 2.88; Score 36.6; DB 1; Length 120;
Best Local Similarity 71.68; Prid. No. 0.3;
Matches 48; Conservative 0; Mismatches 19; Indels 0; Gaps
1
1232 gtaataaagatgtaagaaataagatatattgtagacaatatcttcataaaaaaaaaa 1297
Db 30 GTAATTACTTGCAATATGAGAAATAAACAATTGGCACAATTGTATAAAAAAAAAA 89
Cy 1292 aaaaaa 1298
Db 90 AAAAAA 96

```

RESULT 9
US-08-151-477A-28
; Sequence 28, Application US/08151477A
; Patent No. 5830644
GENERAL INFORMATION:
APPLICANT: Michael D. West
APPLICANT: Jerry W. Shay
APPLICANT: Woodring E. Wright
APPLICANT: Elizabeth Blackburn
APPLICANT: Nam Woo Kim
APPLICANT: Calvin B. Harley
APPLICANT: Scott L. Weinrich
APPLICANT: Catherine Strahl
APPLICANT: Michael J. McEachern
APPLICANT: Homayoun Vaziri
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF
TITLE OF INVENTION: CONDITIONS RELATED TO
TITLE OF INVENTION: LENGTH AND/OR TELOMERASE ACTIVITY
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/151,477A
FILING DATE: No. 5830644ember 12, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/038,766
FILING DATE: March 24, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 202/189
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 120
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-151-477A-28

Query Match 2.8%; Score 36.6; DB 2; Length 120;
Best Local Similarity 71.6%; Pred. No. 0.3;
Matches 48; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 1232 gtaataaagtgatgcaagaataaagatattagacatatattcataaaaaa 1291
DB 30 gtaattactgcattgacgaataaacaattggcacattgtttaaaaa 89
OY 1292 aaaaaa 1298
DB 90 AAAAAA 96

RESULT 10
US-08-819-867-58
; Sequence 58, Application US/08819867
; Patent No. 6007989

GENERAL INFORMATION:
APPLICANT: Michael D. West
APPLICANT: Calvin B. Harley
APPLICANT: Scott L. Weinrich
APPLICANT: Catherine M. Strahl
APPLICANT: Michael J. McEachern
APPLICANT: Jerry Shay
APPLICANT: Woodring E. Wright
APPLICANT: Elizabeth H. Blackburn
APPLICANT: Nam Woo Kim
APPLICANT: Homayoun Vaziri
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF
TITLE OF INVENTION: CONDITIONS RELATED TO
TITLE OF INVENTION: TELOMERASE LENGTH AND/OR
TITLE OF INVENTION: TELOMERASE ACTIVITY
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,867
FILING DATE: March 14, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/153,051
FILING DATE: No. 6007989ember 12, 1993
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Chambers, Daniel M.
REGISTRATION NUMBER: 34,561
REFERENCE/DOCKET NUMBER: 224/232
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-819-867-58

Query Match 2.8%; Score 36.6; DB 3; Length 120;
Best Local Similarity 71.6%; Pred. No. 0.3;
Matches 48; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 1232 gtaataaagtgatgcaagaataaagatattagacatatattcataaaaaa 1291
DB 30 gtaattactgcattgacgaataaacaattggcacattgtttaaaaa 89
OY 1292 aaaaaa 1298
DB 90 AAAAAA 96

RESULT 11
US-09-120-365-88
; Sequence 88, Application US/09120365
; Patent No. 6103514

```

: GENERAL INFORMATION:
: APPLICANT: Natori, Shunji
: TITLE OF INVENTION: NEW PROTEASE
: FILE REFERENCE: 32290-144749
: CURRENT APPLICATION NUMBER: US/09/120,365
: CURRENT FILING DATE: 1998-07-22
: EARLIER APPLICATION NUMBER: JP 9-333 474
: EARLIER FILING DATE: 1997-11-18
: NUMBER OF SEQ ID NOS: 101
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 88
: LENGTH: 1916
: TYPE: DNA
: ORGANISM: Sarcophaga peregrina
: US-09-120-365-88

Query Match          2.8%; Score 35.8; DB 3; Length 1916;
Best Local Similarity 53.6%; Pred. No. 1.6;
Matches 96; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 1116 ataatacttcatcttcctgcgcacaaagccagagagcagcgtgtaaaagtttg 1175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1739 ataattgcggtgtatagcactgcccactatgtagagatgacatattgattctcg 1798
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1176 ttattagttggttttgcgcgttggtggtggttgcgtggttgcgttgcgttaa 1235
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1799 cttaagccttataaaagacttaatttcatttaatttccttaattttttttta 1858
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1236 taaaagtgtgcaagaataagatattagtaacaatttccttaaaaaa 1294
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1859 t-aagttgtctactcaataataagaagaataaattgttttaataaaaaa 1916
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

```

```

: US-09-061-768A-3
: Sequence 3, Application US/09061768A
: Patent No. 620437
: GENERAL INFORMATION:
: APPLICANT: BRASH, ALAN R.
: APPLICANT: BOEGLIN, WILLIAM E.
: APPLICANT: JISAKA, MITSUO
: TITLE OF INVENTION: LIPOXYGENASE PROTEINS AND NUCLEIC ACIDS
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARLES A. TAYLOR, JR.
: STREET: SUITE 1400, UNIVERSITY TOWER, 3100 TOWER BOULEVARD
: CITY: DURHAM
: STATE: NORTH CAROLINA
: COUNTRY: USA
: ZIP: 27707
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB storage
: COMPUTER: IBM PC/XT/AT compatible
: OPERATING SYSTEM: Windows 3.1
: SOFTWARE: WORD PERFECT 6.1 and ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/061,768A
: FILING DATE: APRIL 16, 1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA: NONE
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: ARLES A. TAYLOR, JR.
: REGISTRATION NUMBER: 39,395
: REFERENCE/DOCKET NUMBER: 1242/5
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 493-8000
: TELEFAX: (919) 419-0383
: TELEX:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3205 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: US-09-061-768A-3

Query Match          2.7%; Score 35.2; DB 4; Length 3205;
Best Local Similarity 52.8%; Pred. No. 2.8;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1155 agtcaagcgttaaaagttttttagttgttttgcgttggtggttgcgtt 1214
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3061 acitttgctgctgacggcgttttttttgcgtttggttgcgttggttgcgtt 3120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1215 attgctgtttgtattcgttaataaagtgatgcaagaataagatattagtaacat 1274
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3121 tttttgattttgtttttctgttcacactcagacgttcatttaattgcatatga 3180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1275 ttcatataaaaaa 1298
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3181 gttcatttcacaaaaa 3204
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
: US-08-330-108-16
: Sequence 16, Application US/08330108
: Patent No. 5795752
: GENERAL INFORMATION:
: APPLICANT: Smith, Kendall A.
: TITLE OF INVENTION: IL-2-Stimulated Gene
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lathive & Cockfield

```

```

SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10087
FILING DATE: 19921118
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/796,066
FILING DATE: 20-NOVEMBER-1991
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-028PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-7400
TELEFAX: 617-227-5941
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
LIBRARY:
CLONE: 8D4-T3
PCT-US92-10087-16

Query Match      2.7%   Score 35; DB 5; Length 198;
Best Local Similarity 55.3%; Pred. No. 1;
Matches    68; Conservative     0; Mismatches    55; Indels    0; Gaps    0;

OY 1176 ttttaagtggttttgcgctgggggcttcctcgagtattgtcgttttgtatcgttaa 1235
DB 52 ttittcttttccctctgttttttatTTTGGATTAAGTCGCTTGCCCTTATTGTAAACGTGTA 111
OY 1236 taagaagtcgtgcacaagaataagacatttagtacaaatcttcataaaaaaaaaaaaa 1295
DB 112 TAAATATATATTATATATAATATTTMAAAGAATAATGTTCAGAATAAAAAAAAAAAAAA 171
OY 1296 aaa 1298
DB 172 AAA 174

Search completed: July 27, 2002, 04:40:45
Job time: 4466 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 27, 2002, 03:21:59 / Search time 1730.1 Seconds
(without alignments)
10126.032 Million cell updates/sec

Title: US-09-971-020-2

Perfect score: 1 agcagcgcgcacatcgcattgt.....ataaaaaaaaaaaaaaa 1298

Sequence: 1 IDENTITY NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estov:*
7: em_estro:*
8: em_estro:*
9: em_estro:*
10: em_estro:*
11: em_estro:*
12: em_estro:*
13: em_estro:*
14: em_estro:*
15: em_estro:*
16: em_estro:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198.8	15.3	848	10	BS586603 EST488371
2	161.8	12.5	676	10	BI217170 NR011D12
3	146.2	11.3	542	9	BE202402
4	145.6	11.2	637	10	BE202402
5	145	11.2	495	9	AL368351
6	136.2	10.5	604	10	BE004423
7	129	9.9	539	10	BE004423
8	129	9.9	539	10	BE004423
9	129	9.9	539	10	BE004423
10	126	9.7	716	10	BS585368
11	125	9.6	711	10	BE202392
12	124.8	9.6	474	9	BE202392
13	117.2	9.0	723	10	BS585571
14	109.8	8.5	482	10	BI205540
15	107.8	8.3	455	10	BI210946
16	107.8	8.3	455	10	BI210946
17	107.8	8.3	500	10	BI973953

18	106.2	8.2	637	9	AM217769
19	106	8.2	451	9	AL369180
20	104.6	8.1	466	10	BS586994
21	104.6	8.1	466	10	BS586994
22	104.2	8.0	481	10	BE789602
23	104.2	8.0	481	10	BE789602
24	104	8.0	495	10	BE004501
25	104	8.0	495	10	BE004501
26	104	8.0	535	9	AM73568
27	103.6	8.0	568	9	AM584129
28	103.2	8.0	466	10	BE042646
29	102.6	7.9	483	10	BE598091
30	102.2	7.9	561	9	AM217770
31	102.2	7.9	561	9	AM217770
32	102.2	7.9	561	9	AM217770
33	102.2	7.9	561	9	AM217770
34	102.2	7.9	561	9	AM217770
35	101.8	7.8	568	10	BM301977
36	100.8	7.8	702	10	BM413186
37	97.6	7.5	698	10	BE69572
38	97.6	7.5	503	10	BE319252
39	93.8	7.2	588	10	BE097739
40	92.8	7.1	545	9	AM298283
41	92.2	7.0	584	9	BE205019
42	91.4	6.9	561	10	BE462869
43	90.2	6.9	561	10	BE358819
44	89.8	6.8	578	10	BE359581
45	88.8	6.8	578	10	BE359581

ALIGNMENTS

RESULT 1
LOCUS BS586603
DEFINITION EST488371 MHAM Medicago truncatula/Glommus versiforme mixed EST
ACCESSION BS586603
VERSION 1
KEYWORDS
SOURCE Medicago truncatula/Glommus versiforme mixed EST library
ORIGIN Medicago truncatula/Glommus versiforme mixed EST library
REFERENCE 1 (bases 1 to 848)
AUTHORS Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
and Fraser,C.M.
TITLE ESTs from roots of Medicago truncatula after colonization with
Glommus versiforme, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Noble EST name: N385669e TIGR sequence name: MTDCULIRK More
information is available at: <http://www.medicago.org>
Seq primer: Symod (CTA gaa gta gaa CC).
Location/Qualifiers
1. 848
/organism="Medicago truncatula/Glommus versiforme mixed EST
library"
/cultivar="Medicago truncatula genotype A17"
/db_xref="taxon:119092"
/clone="PMHAM-44A21"
/tissue="roots"
/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
post-inoculation with Glommus versiforme. The library was
made from a mixture of RNA from each of these stages."
/lab_host="E. coli strain XL0LR"

from the recombinant lambda-zap phage using Ex-assis-
t helper phage and propagated in XL0L cells."

Query Match	11.38;	Score 146.2;	DB 9;	Length 542;
Best Local Similarity	59.20;	Pred. No. 1.2e-17;		
Matches 313; Conservative	0;	Mismatches 198;	Indels 18;	Gaps 3

[illegible]

from the recombinant lambda-zap phage using Ex-assist
helper phage and propagated in XL0R cells."

BASE COUNT	163 a	105 c	104 g	170 t
Query Match	11.3%	Score 146.2;	DB 9;	Length 542;
Best Local Similarity	59.28;	Pred. No. 1.2e-17;		
Matches 313;	Conservative 0;	Mismatches 198;	Indels 18;	Gaps 3;

OY	13	tgcattgctccgatatagtatgaagtcctccaagaagtctcgtcatatgaatgaaggtagg	72
Db	15	TGCTAGACGCCCTTTTACCATTGCACATGGCAGCTTCACATCATGTACTGACGTGAGCG	74
OY	73	cgalacaagctacggccaagaatgcalactc---acaatcttgctcttgcgaagtgtaaacc	129
Db	75	AGATTAACAAGCTATCCAAATAAATACGCCAACATTTCAAGAATGTGATGCTGACGCGAAMACA	134
OY	130	tttcccttgacaatatgcatcgagaatttgttcggcgcaaccttgcaccaatcacaaaagt	189
Db	135	TAAACTTGGAAGAGATTAATVAGSGCTGATTGTGACACACTTTTCCAAT-----TG	185
OY	190	caltaaatgttcggagattltyggatlgcgtctcttgacaaccaaacacettttaacagtgcgga	249
Db	186	TTTTGAAMAAGGCGCTACTTAGAGTTGTTCATCAGGTGCACAAAAGCACCTTTGTAGCATCAA	245
OY	250	catctgcgcaaatlttacaaaaatltygccagaagaagaagaatgaatttgaacgttcccac	309
Db	246	TATCAATAAACACTATTGATGCTGTGAGTCMAAAAAATFAGACATGAGATCA-----CCCAT	299
OY	310	catccggaattttctcgaatgatcttttccaaaatgatctaattcggtttcaaagtgtct	369
Db	300	GTTTCAGTGTTTTCCCATCAATGACCTATTGGAATGACGTTCMAATCCACACTTTTAAGTTACT	359
OY	370	gccaaagcttccacgcgcaaatcgagaagaagaatgagcgaagaatagatgagtcgtgccaat	429
Db	360	ACCTGATTTTATAAAAAAGACTACACMAAAGAAAGAACAAATAATTGATGCATGTTCTT	419
OY	430	aagcgaagaagcgtgcgtcttctcagaagcgagactcttccccgagagatccatgatctttt	489
Db	420	TATGTGAACACTGGCTTTTATATGGAGACTCTTCCCGCACAATTCATTCACACTTTT	479
OY	490	gcaccttgttatcagtlgttcatgtgtlatcctaagtttcccacgggttg	538
Db	480	TCATTCTCTCTATGACTTACACTGGCTTTCCACAGACTCCAGATGCATTG	528

RESULT	4				
LOCUS	BF004424				
DEFINITION	ESTJ32322 KYI Medicago truncatula DNA clone pKVL1-17L3,	mRNA	linear	EST	06-OCT-2000
ACCESSION	BF004424				
VERSION	BF004424				
KEYWORDS	sequence.				
SOURCE	Medicago truncatula				
ORGANISM	bareil medic.				
REFERENCE	1 (bases 1 to 637)				
AUTHORS	Vandenbosch K., Endre-G., Hur J., Moore J., Beremand P., Ellis L., Town C.D., Bowman C.L., Craven M.B., Hansen T.S., Holt I.E. and Fraser C.M.				
JOURNAL	ESRs from roots of Medicago truncatula 24 hours after inoculation with Sinorhizobium meliloti				
COMMENT	Unpublished (1999) Contact: Vandenbosch K. Department of Biology Texas A&M University College Station, TX 77843-3250, USA Tel.: 409 845 7707				

was directionally ligated into the Unizap XR vector from
Stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing CDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

BASE COUNT 159 a 128 c 76 g 176 t

Query Match 9.9%; Score 129; DB 10; Length 539;
Best Local Similarity 55.1%; Pred. No. 1.9e-14;
Matches 252; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

615 caatttgatcaattacgaagaatttaccacatttcaaggaattcagaagaagt 674
|||||
Db 539 CATACCTTGAACAAATTCACACAGATTCATATTATTATTAAGATCAGCTGCTGAC 480
675 tgtttccagtgccgaatgctccttaccctgcattgttaagtagaattcgaagaac 724
|||||
Db 479 TGGTTCCTAGTGGTGCATGCTTAACTATTTGGCAGAGATGACCAAAATGACTTA 420
673 cgaatcccttagacttacttgcacgaataaagacttgaattgaaggaattcttg 794
|||||
Db 419 TGAATGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 360
675 aagcaagaagaatgagatgattcaatttccatttaccacattcagaagaagaata 854
|||||
Db 359 AGCAATCAAAATTTGACATTTTACATACCATCTTATGCTGCTGATGATGATGAT 300
655 agtgcatagttgaaggaagagcttcctgcgaattatatacttgaagactttaaggcc 914
|||||
Db 299 GAAAGTGATGAG 240
655 attatgagctgccttctctatcttgaatgattaccagaaagacccatgaacaaata 974
|||||
Db 239 ATTGGGTGAAAAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 180
655 aagcaagatattgagcatcaattatagacagtttcaagaaacccatcgcgaagcatt 1034
|||||
Db 179 GGGCTGAGGCTGACCAAGATTATTAAGAGCTGTGACAAACCAATTTGAAGTACAGT 120
655 ttgagaagctatttgccttacttaccacagagct 1071
|||||
Db 119 TTGGAGAGAAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 83

RESULT 8
AM774361/c 685 bp mRNA linear EST 07-SEP-2000
LOCUS EST33512 KV3 Medicago truncatula cDNA clone PKV3-22G18, mRNA
DEFINITION sequence.
ACCESSION AM774361
VERSION AM774361.1 GI:7718276
KEYWORDS EST.
SOURCE
ORGANISM Medicago truncatula
barrel medic.

REFERENCE
AUTHORS Vandenbosch, K., Hunt, J., Moore, J., Beremand, P., Peng, H., Ellis, L.,
Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S., Holt, I. E. and
Fraser, C. M.
ESTs from roots of Medicago truncatula after Rhizobium inoculation
unpublished (1999)
CONTACT: Vandenbosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu

TEXAS A&M EST name: T257780E
TIGR sequence name: MREA457K
More information is available at:
<http://cchr.tamu.edu/medicago>
Seq primer: SKMD (CTA GAT GAT CC).
Location/Qualifiers

FEATURES
source
1. 685
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="PKV3-22G18"
/clone_1db="KV3"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK-; Site, 1: EcoRI; Site, 2:
XhoI; CDNA was prepared from polyA+ enriched RNA. The CDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing CDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

BASE COUNT 228 a 148 c 91 g 218 t

Query Match 9.9%; Score 129; DB 9; Length 685;
Best Local Similarity 55.1%; Pred. No. 1.7e-14;
Matches 252; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

615 caatttgatcaattacgaagaatttaccacatttcaaggaattcagaagaagt 674
|||||
Db 685 CATACCTTGAACAAATTCACACAGATTCATATTATTATTAAGATCAGCTGCTGAC 626
675 tgtttccagtgccgaatgctccttaccctgcattgttaagtagaattcgaagaac 734
|||||
Db 625 TGGTTCCTAGTGGTGCATGCTTAACTATTTGGCAGAGATGACCAAAATGACTTA 566
673 cgaatcccttagacttacttgcacgaataaagacttgaattgaaggaattcttg 794
|||||
Db 565 TGAATGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 506
675 aagcaagaagaatgagatgattcaatttccatttaccacattcagaagaagaata 854
|||||
Db 505 AGCAATCAAAATTTGACATTTTACATACCATCTTATGCTGCTGATGATGATGAT 446
655 agtgcatagttgaaggaagagcttcctgcgaattatatacttgaaggaatttgaagcc 914
|||||
Db 445 GAAAGTGATGAG 386
655 attatgagctgccttctctatcttgaatgattaccagaaagacccatgaacaaata 974
|||||
Db 385 ATTGGGTGAAAAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 326
675 aagcaagatattgagcatcaattatagacagtttcaagaaacccatcgcgaagcatt 1034
|||||
Db 325 GGGCTGAGGCTGACCAAGATTATTAAGAGCTGTGACAAACCAATTTGAAGTACAGT 266
655 ttgagaagctatttgccttacttaccacagagct 1071
|||||
Db 265 TTGGAGAGAAATATATGATGATGATGATGATGATGATGATGATGATGATGATG 229

RESULT 9
AL371374/c 499 bp mRNA linear EST 03-AUG-2000
LOCUS AL371374
DEFINITION MCB443607R1 MCB4 Medicago truncatula cDNA clone MCB443607 T1, mRNA
sequence.
ACCESSION AL371374
VERSION AL371374.1 GI:9671127
KEYWORDS EST.
SOURCE
barrel medic.

REFERENCE	TITLE
AUTHORS	JOURNAL
COMMENT	COMMENT
Medicago truncatula Eukaryotic Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophytia; Magnoliophyta; eudicotyledons: core eudicots; Rosidae, eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliée Medicago. 1 (bases 1 to 499) Journé,E.P., Crespeau H., van Tuinen D., Guzy,J., Tallon,O., Nlebel,A., Carreau,V., Chataigner,O., Kahn,D., Glanlinazzi-Pearson ,V. and Gamas,P. Medicago truncatula ESTs from nitrogen-starved roots Unpublished (2000)	Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email : sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr Contact : Pascal Gamas and Etienne-Pascal Journé, Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, CNRS-IRMA, BP 27 31326 Castanet-Tolosan Cedex, France (Email : Mt.estoulouse.irma.fr Website : http://sequeance.toulouse.irma.fr/Mttruncatula.html). Location/Qualifiers 1..499
FEATURES	/organism="Medicago truncatula" /cultivar="Jemalong" /db_xref="taxon:3880" /clone="MGBA43C07" /clone_1lb="MBA" /tissue_type="root tips" /dev_stage="harvested after 3 days of N-starvation" /note="Vector: pBluescript psk; Site 1: EcoRI; Site 2: XhoI; Plants were grown in an aeroponic chamber for 14 days on nitrogen-rich medium followed by 3 days on N-free medium. RNA was extracted from root tips (1-3 cm). cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into uni-zapR vector from stratagene and packaged using Gigapack Gold packaging extracts. plasmids containing cDNA inserts were mass-excisied from phage stocks using Exasit helper phage and propagated in SGR cells. Clone ordering and sequencing was performed by the Centre National de Sequencage (Genoscope, Evry, France)."
BASE COUNT	168 a 96 c 88 g .
ORIGIN	
Query Match:	9.88; Score 127; DB 9; Length 499;
Best Local Similarity	58.88; Pred. No. 4.5e-14;
Matches 284; Conservative 0; Mismatches 180; Indels 19; Gaps 3	
OY	13 tggattgctcgcatgatgaatgtagtgacccaagaagtccctgtgatatgaatgaagtgaagg 72
Dd	468 TCGTAGCTACCCCTTTAGCAAAAGGAGGTGAACAAGTCCTTGCATCATGTAATGAGAGTGAGGG 409
OY	73 cgatacaagctcagcccaagaagt-----catctacaactctgcgccttcgccaagtgtgaac 128
Dd	408 AGATTAAACACTATGCGAGATTAATCCACATTTTTCAGAGAATGATGCTGACGCCGAAC 349
OY	129 ctctcttgaacaagcataagtagaagtgttgcggccaactttgccaaatcataaagt 188
Dd	348 ATTACTTGAAGAAGATTAATGAGGCTGATTAATGATGACACTTTTCCAAN -----T 298
OY	189 gacttaaatlttgagatttgtagatgcgttcttcggccaacaacacctttaaacagtgcgg 248
Dd	297 GTTTGAATGCTGACTTACCTTACGTTTCATCATGAGTCCAAAATGCACTTTTGATGATCAA 238
OY	246 acatttgcgaagaattttcacaaagtgtgcgcaagaagaagaagaatgaattgaacgtccca 308
Dd	237 ATATCATTAACACTATATATCTGTGATCAAAAAATGTGATCAATGATCA-----CCCA 184
OY	309 ccattcaagattttctggaagatctttccaaaatgaattcaatgcagtttcaagtgc 368
Dd	183 TGTTTGAATTTTTCTCATATACCTATTATTGGAATGACTCAATPACCACTTTTAAGTTAC 124

```

Oy 369 tgcgaagcttcaccgcaaacctgcggaaagaataatgacgcaagatagatcgcgctaa 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 TACGATATTATTAATAAGACATCAAGAGAGAAAGACAAAAGTTTATGTCATGTTTC 64

Oy 429 taagggcaatgcctgcgcctctcttcacggcaagactctcccgaggagatcatgatttt 488
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 TAGAGCAACACCTGGGTCTTTTATGGGACACCTCTCCCGACAAATTCATTACTTTT 4

Oy 489 tgc 491
      |||
Db 3 TTC 1

RESULT 10
BG585368 716 bp mRNA linear EST 11-APR-2001
LOCUS DEFINITION EST487133 MHAM Medicago truncatula/Glomus versiforme mixed EST
ACCESSION Library CDNA clone PMHAM-22N19.5' end, mRNA sequence.
SOURCE BG585368
KEYWORDS BG585368.1 GI:13600432
EST.
REFERENCE Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM Medicago truncatula/Glomus versiforme mixed EST library
AUTHORS Eukaryota; mixed EST libraries.
1 (bases 1 to 716)
Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.T.
and Fraser,C.M.
ESTs from roots of Medicago truncatula after colonization with
Glomus versiforme, 2001
Unpublished (2001)
JOURNAL Contact: Harrison M.J.
COMMENT The Samuel Roberts Noble Foundation.
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5811
Fax: 580-221-7380
Email: mjharrison@snoble.org
Noble Est name: NS80842e TIGR sequence name: MPOC82PK More
information is available at: http://www.medicago.org
Seq primer: Skmod (CTA GAA CTA gtc gat CC).
Location/Qualifiers
1. 716
/organism="Medicago truncatula/Glomus versiforme mixed EST
library"
/cultivar="Medicago truncatula genotype A17"
/db_xref="taxon:119092"
/clone="PMHAM-22N19"
/clone_1lb="PMHAM"
/tissue_type="roots colonized with Glomus versiforme"
/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
post-inoculation with Glomus versiforme. The library was
made from a mixture of RNA from each of these stages."
/lab_host="E. coli strain XLOLR"
/notes="Vector: p Bluescript SK-; Site.1: EcoRI; Site.2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
roots harvested at 10, 17, 22, 31 and 38 days
post-inoculation with Glomus versiforme. The cDNA was
directionally ligated into the Unisap XR vector from
stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XLOLR cells."
BASE COUNT 238 a 114 c 167 g 197 t
ORIGIN
Query Match 9.7%; Score 126; DB 10; Length 716;
Best Local Similarity 55.5%; Freq. Mod. 6.1e-14;
Matches 284; Conservative 0; Mismatches 225; Indels 3; Gaps 2
Oy 563 aaccaaaggagatattactcttcacaaagatgctgcgcgcgtccacagaagcatattg 622
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 AACAAAGGGAATATTTCCTTAACGTAACAGACACCCTCCACGACTGCAAAAACGACTTT 202

```

QY 523 gacacatttaagaagaatttaccacatttcaagagttcaatcgaagaagtgttttca 682
 Db 203 GAACATTTCAACAAGATTTCTCATATTTTTTAAGATCAAGCTTCGCTGAACTGCTCC 262
 QY 683 cgttgccgaatagtcctcttaccgcatctgttaagtagaataatgcgaacgaatccc 742
 Db 263 AGTGGGCGCAATGGCTCTAACACTTATTTGGCAGAGATATGCAAAATGAACCTTATTAAGCA 322
 QY 743 cttagact-tacttgacatgycataaagaacacttgatctgttgaaggaactcttcygaaga 801
 Db 323 TGGGGTTGTCATTTGGCATGGCACCTCAATGACATGGCCGACGTCAAGTTGGTGAGCAATC 382
 QY 802 aaatattgatatttcacatattccattcttaccaccttcagacgaagaagaatagtgcat 861
 Db 383 AAAATTTGGCATTTTAACTACATACATCCATTTGGCCCTGATCGATGAGATTGAAGTAAAAGT 442
 QY 862 agttgaggaagaagaagttcttgcgaaattttatcttcygaagacttttaagcccatctga 921
 Db 443 GATTGAAGAAACAAAGGCTTTTGAATGTTCAAAAGTTTGAACACATTTAGAACAGGATTTGGGT 502
 QY 922 tgcctgcctctctatctgatgatgathaccocgaataagatcccatgaaacaattaaag- ca 979
 Db 503 GAAAACGTCGATCTGATCCATGATGAGTACACTGTTGTTGATGAAAGAACAGGGCCCT 562
 QY 980 gagtatgycgatcatatattagatcaagtttaagaaaccccttcgcgaagcatcttggga 1039
 Db 563 GAAGGTGAGCAAGTATTAAAGCTGTGTGCACAAACCAATTTTGAAGTCAGAGATTGGGA 622
 QY 1040 gaagctattatgctcgtacttattccaagacct 1071
 Db 623 GAAGAAATTAATGATGATGATTAATTAATTCATTAGGTT 654

RESULT 11	
B1927794	711 bp mRNA linear EST 18-OCT-2001
LOCUS	BT54547683 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
DEFINITION	clone CTDB27F11 5' end, mRNA sequence.
ACCSSION	U17370

ACCESSION	B1921194
VERSION	B1927794.1
KEYWORDS	GI:16238117
SOURCE	EST.
tomato.	

ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; eusterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
1 (bases 1 to 711)

AUTHORS
van der Hoeven, R. S., Bezdekides, J. L., Karamchcheva, S. A., Tsai, J.,
Uttarack, T., Van Aken, S., Romaniuk, C. M., Nierman, W., Fraser, C. M.,
Martlin, G. B., Giovannoni, J. J., and Tankley, S. D.
TITLE
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL
Unpublished (2001)
COMMENT
Contact: CUGI

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics Institute
Seq primer: T3.

```

FEATURES
source
  location/Qualifiers
    1. 711
      /organism="Lycopersicon esculentum"
      /cultivar="PA496"
      /db_xref="taxon:4081"
      /clone="COT23P11"
      /clone_lib="tomato flower, 3 - 8 mm buds"
      /tissue_type="flower"
      /dev_stage="3-8mm buds"
      /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Corneil University; sequencing: The

```

Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, 72496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

Query Match	9.68;	Score 125;	DB 10;	Length 711;
Best Local Similarity	53.58;	Pred. No. 9.4e-14;		
Matches 375; Conservative	0;	Mismatches 290;	Indels 36;	Gaps 4;

Oy 27 tatgaatagaagctcccaagaagtcctccatcatgaatagaagtcgaagcgaatacaagctacg 86
 Db 29 TAGAATGAGAGGTGTTGAGATTCTTCACATGAATGAGGAAATGAGACATTTACTATG 88
 Oy 87 ccaagaatgacaccc---tacaatcgtgcctctgcagaagtgaacctcttcctgaacat 143
 Db 89 CAATATATTCCTGGTTCAGAAAAAGGTATCTCATGACAAAGCCAAATGAAGATCAAG 148
 Oy 144 gcatagagaatctctgcggccaactctcccaacatcaacaagtgcataaagtctggg 203
 Db 149 CCAATAGATCTCTATTGTAACCTATTCGCCAATA-----CATATCAATGGCTG 199
 Oy 204 atttggatgaagctcttcggagccaacaacaccttaacagctggaggaactctgccaagta 263
 Db 200 ATTGGGTTGTTCTTCGAGGACAACTTTTATAGTGATGATCAAGACTGTATAGTGCA 259
 Oy 264 ttgacaagttggccaggaagagaagaatgaattagaagctgcacacctgagattctc 323
 Db 260 TTGAAAAAGAACG-----AAAAAACAGATCTACATCCAGCATGTTTATTTTCACT 313
 Oy 324 tgaatgactcttccaaatgatttaattatctggtttcaagttgcgtgcgccaagctctcac 383
 Db 314 TCAATATCTCCCTGGAATGATTTTAAAGCATTTTTCACATGTTTAGGGGAATTTGAC 373
 Oy 384 gcaactcgagaagaagaatgagcgaagaatgatatgctctccaaataagcgaagtctg 443
 Db 374 AAAATTGAAAAACAAATTTGAGAAAGAACTGTGTCATGTTTTTTTAACTGTGCGCTG 433
 Oy 444 gacctctcgaagcagaagctctcccgagagatccaaagatcttttgcactctgtaca 503
 Db 434 GTTCATTTTATCTGACACTTTTCCATCAAAAGTTTGCAATTTGTCTCATCTCTATA 493
 Oy 504 gttcatattgattctcgaagtcgaagtcgaagcaggtgttgatgaattggagattgtgtcaa 563
 Db 494 GTCTTATGGCGCTATCCAACTTCCTAAATTTA-----ATTGAAAGAA 535
 Oy 564 acaaaaggagattactctctccaagaagctgtgcgcgcgcgcgaagaggaattgtg 623
 Db 536 ACAAGGGAATTTTAAATGCGAAGTACAGTCCACCAAGTGTATATAAAAGATTTTCA 595
 Oy 624 atcaattacagaagaatttacaacactctcagaatctatctgaaagagtggtttcac 683
 Db 596 AGCAATATGCAAAAGATTTTTCATATTTTTCATAATATCGTTCGGAAGATGATGAAG 655
 Oy 684 gtggccgaatgctcctactcgtcatctgttaagtatgaa 724
 Db 656 GTGGAAAAATGGTATTTACATTTTTTGAAGAAAGCAAGTGA 696

RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BE202392	474 bp	BE202392	BE202392	EST	barrel medic.	Medicago truncatula	
BE202392	linear	EST928241 KVI	BE202392	EST	Medicago truncatula	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
	CDNA clone pXVI.265	sequence.	BE202392.1	GI:8745693			

REFERENCE 1 (bases 1 to 474)
 Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,
 Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and
 Fraser, C.M.

TITLE
 ESTs from roots of Medicago truncatula 24 hours after inoculation
 with Sinorhizobium meliloti

JOURNAL
 Unpublished (1999)
 CONTACT: Vandenbosch K
 Department of Biology
 Texas A&M University
 College Station, TX 77843-3258, USA
 Tel: 409 845 7707
 Fax: 409 845 2891
 Email: kate@mail.bio.tamu.edu
 Texas A&M University: T261035
 TIGR sequence name: MTIA397K
 More information is available at:
 http://chrystle.tamu.edu/medicago
 Seq primer: Skmod (CTA gaa gta gat cc)
 Location/Qualifiers

FEATURES
 source
 1..474
 /organism="Medicago truncatula"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="PKV1-263"
 /clone_1b="KV1"
 /tissue_type="Seedling roots"
 /dev_stage="24 hours post-inoculation with Sinorhizobium
 meliloti"
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
 was directionally ligated into the Unipap XR vector from
 Strategene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-zap phage using Ex-assist
 helper phage and propagated in XL0R cells."
 BASE COUNT 149 a 87 c 93 g 145 t
 ORIGIN

Query Match 9.6%; Score 124.8; DB 9; Length 474;
 Best Local Similarity 58.9%; Pred. No. 1,2e-13;
 Matches 279; Conservative 0; Mismatches 177; Indels 18; Gaps 3;
 13 tcgattctctgcacatgagtgagtcacagaagtcctgcacatgagtgagtgag 72
 15 tgcgtacgacacccctttagcactgacatgacacacgctcctcaccatgagtgag 74
 73 cgatacagctacgacagaatgacatcct--acaatctgctcttgccaaggtgaacc 129
 75 AGATACAAAGCTATGCAATAATCACTCCACATTTGAGAAATGCTGACGCGCAAAACA 134
 130 ttccctgacacatgacacagagatgtgtgagcgacacacacacacagaatg 189
 135 TATACTTGAAAGAGTATATGAGGCGTATGTGTGACACTTTCCAAAT-----TG 185
 190 catlaaagctcgcagattgagtcgctctgcagacacacacacacacacacagtcg 249
 186 TTGTAAGAGGCTAGCTAGCTTTTCATCAGTCCAAATGCACTTTGGTAGAGTAA 245
 250 catgtgcagagatgacacagatgtgcagagagagagagagatgaagtcacac 309
 246 TATCATTAACACTATATGCTGTGAGTCAAAATTTGAGTGTAGTGA-----CCCA 289
 310 catcagatcttctgagatcttctcacaagatgacatcggttccaagtgtc 369
 300 GTTTCAGTTTTCCTCAATGACCTATTGGAAATGACTTACCACTTTTAAGTTACT 359
 370 gcccaagctctacgcgaactcgagaagaagaatgagcgaagatagatcgctgcaat 429

DB 360 ACCGATTTTATTAAGACACTACAGAGAGAGCAAAAGTTAGTCATGTTCTT 419
 QY 430 aagcgacatgcctgcctcttctacgagagatcttcccgagaggtcacaaga 483
 DB 420 TAGTGGACACCTGGGCTTTTATGAGGAGACTTCCCGCAATTCATCA 473

RESULT 13

LOCUS
 DEFINITION
 EST487335 MHAM Medicago truncatula/Gloms versiforme mixed EST
 library cDNA clone PMHAM-30124 5' end, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Unpublished (2001)
 Contact: Harrison M.J.
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 Noble EST name: N381045e TIGR sequence name: MTDC607K More
 information is available at: http://www.medicago.org
 Seq primer: Skmod (CTA gaa gta gat cc)
 Location/Qualifiers

FEATURES
 source
 1..723
 /organism="Medicago truncatula/Gloms versiforme mixed EST
 library"
 /cultivar="Medicago truncatula genotype A17"
 /db_xref="taxon:119092"
 /clone="PMHAM-30124"
 /clone_1b="MHAM"
 /tissue_type="Roots colonized with Gloms versiforme"
 /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
 post-inoculation with Gloms versiforme. The library was
 made from a mixture of RNA from each of these stages."
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 roots harvested at 10, 17, 22, 31 and 38 days
 post-inoculation with Gloms versiforme. The cDNA was
 directionally ligated into the Unipap XR vector from
 Strategene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-zap phage using Ex-assist
 helper phage and propagated in XL0R cells."
 BASE COUNT 213 a 136 c 157 g 217 t
 ORIGIN

Query Match 9.0%; Score 117.2; DB 10; Length 723;
 Best Local Similarity 52.9%; Pred. No. 2.6e-12;
 Matches 363; Conservative 0; Mismatches 293; Indels 30; Gaps 4;
 51 tgcataatgaagtgagagtgagagtgacagctcgcagagatgacatcctcagtcg 110
 DB 3 TCCACATGAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
 QY 111 cctctgcagagtgaaaccccttctcgaacatgacatgagatgtgtggcgcaat 170
 DB 61 AAGGAGATTTCTTAAACAACATATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 116
 QY 171 tgcacacatacaacaagtgatcaagtgatgagatgtgagtcgcttcgacaca 230


```

Db 117 CCACGACACCTCCGAGAGCCCTTGCAATGAGATCGGTGGTTCTTGAGACAAACA 176
QY cactitaaagctgagcagcatgtgcaagatgacaaagtgtgacgaagaaga 290
Db 177 CTTGTCTGTGAGTGCAGAAATCATCATGTGTGGAGAAAGCTTGCCACAAATGAAAT 236
QY 291 atgattagaagctgccaccatcagattttctcgaatgacatttcccaaatattca 350
Db 237 ATGCATC-----TCCGCAATACAAAGATCTTTTAAACATCTTTCAGGAATGCTTCA 290
QY 351 attcggttttcaagttgtcgcgaagcttttaccgcaacctgcgaagaagaatgagcga 410
Db 291 ACAACATTTTAAAGTACACTGTGATGACTTTAAACATAACTACTGTATGAAATCAAACTG 350
QY 411 agataagatctggtcctaaagaacgaatgctgtccttttctcagcagactcttcccg 470
Db 351 AAATGACCTTGTCTACTTCTTTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
QY 471 agagatcattgattttttgacactcctgttaccagtgatcattgttattcaggttcca 530
Db 411 ATCGAAGTCTGCAATTCGTCACCTGCTTACAGCTCTTACAGCTGCTGCTGCTGCTG 470
QY 531 ggcgtttgtgattgattgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 590
Db 471 AAGGT-----ATAGATTAACAATAAGGTAACATTTACATCAACGACA 512
QY 591 gatgcgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 650
Db 513 CAATGCTCCCTCAATGTCGCAAGGCTTACTACGACATTTCAAGAGACTGTCGATTT 572
QY 651 ttccaagatcattcgaagaagtggttttaccgtgcgaagtcgccttccctcagcatt 710
Db 513 TTTCAGAGTGTCCCGCAAAAGAAATGTTGTAAGGGGTGTCGATGTTTCAACGATGTTG 632
QY 711 gtaagaagtagtattcgagcgaacg 736
Db 633 GAAAGAAAATGAAGATCCATGTGACG 658

RESULT 14
LOCUS B1205540 482 bp mRNA linear EST 11-JUL-2001
DEFINITION EST533580 cTOS Lycopersicon esculentum cDNA clone cTOS8B4 5' end,
mRNA sequence.
ACCESSION B1205540
VERSION B1205540.1 GI:14683264
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 482)
REFERENCE 1 (bases 1 to 482)
AUTHORS van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R.,
Ronnig,C. and Tanksley,S.
TITLE Generation of ESTs from Tomato Suspension Cultures
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1. 482
/organism="Lycopersicon esculentum"
/cultivar="TA496, E6203"
/db_xref="taxon:4081"
/clone="cTOS8B4"
/clone_1db="cTOS"
/tissue_type="suspension cultures"
/lab_host="SOLR"
FEATURES
Source

```

```

BASE COUNT 153 a 81 c 113 g 135 t
ORIGIN
Query Match 8.5%; Score 109.8; DB 10; Length 482;
Best Local Similarity 57.6%; Fred. No. 7.3e-11;
Matches 223; Conservative 0; Mismatches 152; Indels 12; Gaps 1;
QY 551 gggattgtgcaacaaggagatttactcttccaagatgtcgtccgcgcgcg 610
Db 62 GGAATGTGAATTAACAAAGGACTATTATGATCAAGTCAAGCCATCAAGCGTCTTC 121
QY 611 aagcatatttgatcatttcaatcgaagaatttaccacatttcaagattcattcgaa 670
Db 122 AAAGCATATTCGCAAGCAATATAAAGAGATTTCGCGACTTTTCTCAAGTATCGTCAAGA 181
QY 671 gaggttttcagctgacgacgaatgctccttaccctgatttgaagtagatgattcgac 730
Db 182 GAATTGTGTAAGGTGACGATGATGATTAAGCATGCTGTAAGAAAGAAATGACATCAC 241
QY 731 gaaccgaa-----tcccttagacttacttgacatgacgacataaagcattgatt 778
Db 242 TTAAGCAATGTGTGCGCTTTATGTTGAGCCTTTGGCCATCGCGCTCAAAAGACTGGTT 301
QY 779 gttggagactcttggaggaagaagaattgattgattcattcattcattcattcattc 838
Db 302 ACAGAGGATGATGATGAAGAAAGAAATGATGCTTCACGTTCCACAGTACGTCACA 361
QY 839 tcaagagaagaagaatgacatattgagtgaggaagtggttcttgaggaatttattc 898
Db 362 TCTCCAGCAAAATACATATGATGTTGAAGAGAGATCTTCACTATTGACCTTTTG 421
QY 899 gaggctttaaggccattatgatct 925
Db 422 AGAATCTTAAGAACCAACGATGATCT 448

RESULT 15
LOCUS B1210946 678 bp mRNA linear EST 11-JUL-2001
DEFINITION EST528986 cTOS Lycopersicon esculentum cDNA clone cTOS24D12 5' end,
mRNA sequence.
ACCESSION B1210946
VERSION B1210946.1 GI:14688670
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 678)
REFERENCE 1 (bases 1 to 678)
AUTHORS van der Hoeven,R., Sun,H., Bezzerides,J., Cho,J., Utterback,R.,
Ronnig,C. and Tanksley,S.
TITLE Generation of ESTs from Tomato Suspension Cultures
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1. 678
/organism="Lycopersicon esculentum"
/cultivar="TA496, E6203"
FEATURES
Source

```

```
/db_xref="taxon:4081"
/clone="CTOS24D12"
/clone_lib="CTOS"
/issue_type="suspension cultures"
/lab_host="SOIR"
/notes="Vector: p Bluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Suspension cultures of L. esculentum E6203 were grown
in Murashige and Skoog based medium, supplemented with 15%
coconut milk (filter sterilized and added after
autoclaving), 2% saccharose, and 1mg/ml 2,4D (pH5.8).
Fresh medium was added every 7 days, and cultures were
grown at 25 C, with 12hrs of light and continuous
shaking."
```

BASE COUNT 217 a 102 c 149 g 210 t

ORIGIN

Query Match 8.3%; Score 108.2; DB 10; Length 678;
Best Local Similarity 57.4%; Pred. No. 1.3e-10;
Matches 222; Conservative 0; Mismatches 153; Indels 12; Gaps 1;

```
OY 551 ggagttgtgcaacaaggaggtatttactcttccaaaggatgtcgcgcgtccag 610
DB 62 GGATTTGAGAAATAACAAAGGACTTTATTCATCAAGTACAGCCATCAAGCCTCTC 121
OY 611 aagcatattgtatcaatttcgaagaatttcacacatttcaagattcattcgaa 670
DB 122 AAAGCATTTCCAAAGATATTAAGAAATTTCCGACATTTCTCAAGTATCGTTCAAGA 181
OY 671 gagttgttccagctgcccgaatgtctccttactctgcaattgttaagtagatgaatcgac 730
DB 182 GAATTTGTAAGGTGACGTATGATTTACCATGCTGTGAAGAAATAGCATCTAC 241
OY 731 gaaccgaa-----tccctagacttacttgacatgycataaagcattgatt 778
DB 242 TTAAGCAATGTGTGCTTTATGTTGAGACCTTTGGCCATCGCGCTCAAGAGACTTGTT 301
OY 779 gttagggacttctggaagaagaatttgatagattcaataatccattcttaccact 838
DB 302 ACAGAGGATCGATAGAGAGAGAAATGATTCGTTCAACGTTCACAGTACCTCCA 361
OY 839 tcagcaagaagaatgaagtagttagaagaagaagttcttcggaatttataatctg 898
DB 362 TCTCCACAGCAAAATACGATGATGATTAAGAAAGAGGATCTTCACTATGTGACCTTTTG 421
OY 899 gagactttaaggccattatagct 925
DB 422 AGAATTTAGACACACAGATGATTC 448
```

Search completed: July 27, 2002, 03:57:08
Job time: 2109 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2002, 04:39:44 ; Search time 50.23 seconds
(without alignments)

835,873 Million cell updates/sec

Title: US-09-971-020-1

Perfect score: 1972

Sequence: 1 MEQEVTLHNEGEGDTSYAK.....KGCYNLTLSAKRKPSDV 378

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A.Geneseq.032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	649.5	32.9	364.22	AA67769
2	639	32.4	389.21	AA622189
3	638	32.4	405.21	AA622188
4	628	31.8	386.21	AA622180
5	614	31.1	356.22	AA645739
6	611.5	31.0	379.21	AA642686
7	611.5	31.0	379.21	AA642687
8	609.5	30.9	349.21	AA637439
9	609.5	30.9	349.21	AA637440
10	586	29.7	327.21	AA637440
11	586	29.7	327.21	AA637441

12	432	21.9	374.21	AA606670	Arabidopsis thalia
13	432	21.9	374.21	AA606653	Arabidopsis thalia
14	432	21.9	386.21	AA606659	Arabidopsis thalia
15	432	21.9	386.21	AA606652	Arabidopsis thalia
16	432	21.9	387.21	AA606651	Arabidopsis thalia
17	432	21.9	389.21	AA606658	Arabidopsis thalia
18	395.5	20.1	343.21	AA606664	Arabidopsis thalia
19	395.5	20.1	348.21	AA606663	Arabidopsis thalia
20	393	19.9	345.21	AA613021	Arabidopsis thalia
21	393	19.9	345.21	AA613020	Arabidopsis thalia
22	367.5	18.6	318.21	AA613022	Arabidopsis thalia
23	367.5	18.6	318.21	AA606665	Arabidopsis thalia
24	253.5	15.4	255.21	AA622998	Arabidopsis thalia
25	253.5	15.4	255.21	AA622997	Arabidopsis thalia
26	133	12.9	187.21	AA614384	Arabidopsis thalia
27	129	12.9	154.21	AA614383	Arabidopsis thalia
28	129	12.9	112.21	AA614382	Human protein sequ
29	106.5	5.4	585.22	AA695130	Human RNA helicase
30	106.5	5.4	585.22	AA695130	Human RNA helicase
31	106.5	5.4	778.22	AA695130	Human RNA helicase
32	103.5	5.2	946.21	AA618208	Human protein secret
33	103.5	5.1	997.22	AA618208	Human protein secret
34	100.5	5.1	556.22	AA694696	Human protein secret
35	100.5	5.1	2144.22	AA685029	Human protein secret
36	98	5.0	2185.12	AA612141	Human protein secret
37	97	4.9	960.22	AA69372	Human protein secret
38	96.5	4.9	960.22	AA69372	Human protein secret
39	96	4.8	865.22	AA655771	Human protein secret
40	94.5	4.8	865.22	AA655771	Human protein secret
41	94	4.8	648.20	AA661198	Human protein secret
42	94	4.8	648.20	AA661198	Human protein secret
43	94	4.8	648.20	AA661198	Human protein secret
44	94	4.8	648.20	AA661198	Human protein secret
45	94	4.8	648.20	AA661198	Human protein secret

ALIGNMENTS

RESULT 1	AA67769	standard; Protein; 364 AA.
ID	AA67769	standard; Protein; 364 AA.
AC	AA67769	
DT	11-JUN-2001	(first entry)
DE	Snadragon benzole acid carboxyl methyltransferase (BAMT) protein.	
KW	Snadragon; benzole acid carboxyl methyltransferase; BAMT; floral scent; methyl benzole; transgenic plant; ss	
OS	Antirrhinum majus.	
PN	WO200118183-A1.	
PD	15-MAR-2001.	
PF	01-SEP-2000; 2000MO-US24362.	
PR	03-SEP-1999; 99US-0152393.	
PI	(PUBD) PURDUE RES FOUND.	
DR	Doudareva N, Murfitt LM, Mann CJ.	
DR	WPI: 2001-244568/25.	
DR	N-PSDB: AA680136.	
PT	Novel purified benzole acid carboxyl methyltransferase protein useful for production of floral scent compounds and biosynthesis	
XX		

PS Claim 3; Page 85-88; 94pp; English.

XX The present sequence represents a Snadragon benzoic acid carboxyl

CC methyltransferase (BMT) protein. The protein is a

CC S-adenosyl-L-methionine:benzoic acid carboxyltransferase, and

CC the enzyme is used in the formation of floral scent compounds such as

CC methyl benzoate. The enzyme catalyses the transfer of a methyl group

CC donor, such as S-adenosyl methionine (SAM) for benzoic acid to form

CC methyl benzoate. BMT proteins are useful for the production of floral

CC scent compounds and biosynthesis. BMT polynucleotides are useful for

CC producing transgenic plants.

SQ Sequence 364 AA;

Query Match 32.9%; Score 649.5; DB:22; Length 364;

Best Local Similarity 37.5%; Pred. No. 8.8e-57;

Matches 143; Conservative 85; Mismatches 122; Indels 31; Gaps 10;

QY 3 LQEVLMHN-EGEGDTSYAKNAS-YNLAKVRFLEQICIRRLRANLPINKCIRKADLG 60

DB 4 mklldmldagdsyannsglqkymnkslhvldetlkdlgshv-gfprckfknmdng 62

QY 61 CASGPTLITVADIVOSIDKVGQEKNELEPPTIOIFINDFQNDPNSVFKLPSEYRKL 120

DB 63 csgptlittvadiivosidkvgqekneleppioifindfqnpsvfkllpseyrkl 114

QY 121 EKNGKRKISGCLISAMPGSFYGRLPPEESMHLHSCYVHMLSOVPSGLVIELGICANGK 180

DB 115 shen-----gncfvyglpssfygrllpkkelhlayssylnhslsqpegle-----dnarq 165

QY 181 SYSSKSGRPVQKAYIDQFTTKFTFLRHSEKLEPSGRMLCLCICKVDPEPDPPLD- 239

DB 166 nlymateeppevykaykqyrdftstflklrgeelvpggmvlfngr--svcdpskdd 223

QY 240 -----ILDMKAINDLIVGLLEEKLDSPNIPFTPSAEVNCIYEEGSCFIIILTFEA 294

DB 224 laftllaktlvdmvaelvkmddlylnpdpstcleveaalasgsltdrllevrtv 283

QY 295 HIDAFAISDD---YVPRSHQIKAEYASLRSYEPILASHEGALIMPDLPHILAKHA 351

DB 284 cwdaasyldddqddpslfgkqrsqkivadcvratlepmashstlmolltgykakk 343

QY 352 ANYLHMGKGCYNNLIISLARK 372

DB 344 vehlsvensyfsilvalsir 364

RESULT 2

AAAG22189 ID AAG22189 standard; Protein; 389 AA.

XX AC AAG22189;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 25023.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127463.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130899.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132865.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0136029.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136389.

PR 28-MAY-1999; 99US-0136783.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138340.

PR 10-JUN-1999; 99US-0138844.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139464.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139761.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140699.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142309.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142927.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

[illegible]

PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158382
PR	12-OCT-1999	99US-0158369
PR	13-OCT-1999	99US-0159293
PR	13-OCT-1999	99US-0159294
PR	13-OCT-1999	99US-0159295
PR	14-OCT-1999	99US-0159329
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159331
PR	14-OCT-1999	99US-0159637
PR	14-OCT-1999	99US-0159638
PR	18-OCT-1999	99US-0159584
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160767
PR	21-OCT-1999	99US-0160768
PR	21-OCT-1999	99US-0160770
PR	21-OCT-1999	99US-0160785
PR	21-OCT-1999	99US-0160814
PR	22-OCT-1999	99US-0160815
PR	22-OCT-1999	99US-0160981
PR	22-OCT-1999	99US-0160982
PR	22-OCT-1999	99US-0160983
PR	22-OCT-1999	99US-0161004
PR	22-OCT-1999	99US-0161405
PR	22-OCT-1999	99US-0161406
PR	25-OCT-1999	99US-0161359
PR	26-OCT-1999	99US-0161360
PR	26-OCT-1999	99US-0161361
PR	26-OCT-1999	99US-0161392
PR	28-OCT-1999	99US-0161920
PR	28-OCT-1999	99US-0161943
PR	28-OCT-1999	99US-0162142

Query Match	32.4%;	Score 639;	DB 21;	Length 389;
Best Local Similarity	37.8%;	Pred. NO. 1.1e-55;		
Matches 151;	Conservative	79;	Mismatches 129;	Indels 40;
				Gaps 12;

OY		1	MEDQVFLHMGEGDPTSYAKNKNSTYLALAV-KPELEQRIELLRLAPINKCIKAYDL	59
Dd		1	memvrvllhmngseusyskmscagsnlsigrivmdealklmmsh--selaslyidsl	58
OY		60	GCASEPNTLLTVRDIVQISIDKVGOEKKNELERPT--IQILENDLFQDNFSVKLLPSFY	117
Dd		59	gcspsspnlsistnslvtclhncp---dlrdpyvelrvisindlpndtnylcaaspely	114
OY		118	RKLKENGRKIG-----SCLISAMPGSFYGRLFPESMHRFLSCYSVMHLQVP-	166
Dd		115	--dvnnmkcgfgrggsgecfvaapsafygrlfprshlvfhssslhlwlaqvpcr	174
OY		167	-----SGLVELIGANNKSYSKSGCPRPVOKALDOFTFDFTFLRHSEKEFSBGR	220
Dd		172	eekedrtctcdl---emmgkrylstkcpksahkayalqfgcdfvlrlrsesglvpgr	228
OY		221	MLLTGICR-VDESDEN--PDLIDMANIDLVEELAEKLSQFNIPETPSAEVQC	278
Dd		229	vmystlgrttsldpteesocywellaqqlhsmakejileevetidanaprayaseelkm	286
OY		277	IVEEGSGCEILLYEFTRKAHYDAFSIDDY--PVNSHEOIKR--EYVASIIRSVYPILA	332
Dd		289	vleksgsfstirdleisplwdegwsiseesydlairskpaalasgrvsntllravvephle	346
OY		333	SRRGEAIPMDLFRHLKNAKVALKLMCKGCYNNDIISIAR	371
Dd		349	pftgenvmdelferytklvgeyfyysspryalvllslvr	387

RESULT 3

ID AAG22188

AA AAG22188 standard; Protein; 405 AA.

XX AAG22188;

AC DT

CC 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 25022.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126784.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138440.
PR 10-JUN-1999; 99US-0138447.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0148685.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151063.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

```

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154029.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157855.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.48; Score 639; DB 21; Length 405;
Best Local Similarity 37.88; Pred. No. 1.2e-55;
Matches 151; Conservative 79; Mismatches 129; Indels 40; Gaps 12;

OY 1 MELQEVYLNHNEGCDPTSYAKNASTYNLALAKV-KPFLQOCIRELRLANLPINPKCIRKIVADL 59
DB 17 MEMRYVLMHMKNGSTSYAKNASTYNLALAKV-KPFLQOCIRELRLANLPINPKCIRKIVADL 74
OY 60 GCASGPRITLITVRDVIOSIDRVGOEKNELERPT--IOFLINDLQONQNFVFLPLSPFY 117
DB 75 GCSGSPRITLITVRDVIOSIDRVGOEKNELERPT--IOFLINDLQONQNFVFLPLSPFY 130
OY 118 RKLKENGKRTIC-----SCLISAMPGSFTGRLPPEBSMHFLHSCYSVHMTSOYV-- 166
DB 131 ---drvmnhkgisgfyrggacstvsaavpsaifgrilfiprslthvssssllhwlsqypcr 187
OY 167 -----SGLVIEIGICANKGSTYSKSGCRPPVOKAYLADFTFDFTFTFLIRHSHKELFSGR 220
DB 188 eaekedtltaadi---emngkiyisktkapsahkayalqfcdflvflrseelvpqgr 244
OY 221 MLATCRICK--VDFEPEPN--PDLIDMAINDLIVEGLEEKIKDSPNIFFTPSAEVKKC 276
DB 245 mvslfgrislptleeseygwellaqelmsmakeljeekidafnapyvaasaelkm 304

```

```

OY 277 IVEEGSCCELLYEFKHAHYDAFSDDDY--PYRSHQIRKA--EYVASLIRSYEPIIA 332
DB 305 vlekegfsldrtleapldwegssiseesydlatrskpealasgrtrvanltravemple 364
OY 333 SHFEALMPDLFHLAKHAKAVLHMKCKCYNNLTISLAK 371
DB 365 plfigenwmdelferyakiwgefyfyspsryalvllslvr 403

RESULT 4
AAG22190
ID AAG22190 standard; Protein: 386 AA.
XX
AC AAG22190;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25024.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 30-APR-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 14-MAY-1999; 99US-0134218.
XX
PR 14-MAY-1999; 99US-0134219.
XX
PR 14-MAY-1999; 99US-0134221.
XX
PR 14-MAY-1999; 99US-0134370.
XX
PR 18-MAY-1999; 99US-0134768.
XX
PR 19-MAY-1999; 99US-0134841.
XX
PR 20-MAY-1999; 99US-0135124.
XX
PR 21-MAY-1999; 99US-0135253.
XX
PR 24-MAY-1999; 99US-0135629.
XX
PR 25-MAY-1999; 99US-0136021.
XX
PR 27-MAY-1999; 99US-0136392.
XX
PR 28-MAY-1999; 99US-0136782.
XX
PR 01-JUN-1999; 99US-0137222.
XX
PR 03-JUN-1999; 99US-0137528.
XX
PR 04-JUN-1999; 99US-0137502.
XX
PR 07-JUN-1999; 99US-0137724.
XX
PR 08-JUN-1999; 99US-0138094.
XX
PR 10-JUN-1999; 99US-0138540.
XX
PR 10-JUN-1999; 99US-0138847.

```


of caffeine metabolism related compounds.

Sequence 356 AA;

Query Match 31.1%; Score 614; DB 22; Length 356;
Best local similarity 36.2%; Pred. No. 3,2e-53;
Matches 140; Conservative .78; Mismatches 111; Indels 58; Gaps 11.

XX
SO

9 NMEGGSDPSTAKNNSVYMLALAW-KPELOCIFRELLANPNINNCICIVADLGCASGPT 67
|||:::||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 2 mntgceasyqnsftcqvaaagpalenaveclfrdth--lqelnaadlgcaagpt 59

OY 68 LITARDIYQSIDKKGGEKEKNELEFPI--IQIFINDLQNDFNSVFRLPLSPRYKLENG 125

Dd 60 fa-----vstlkrmekkcetlncqldeqyldlnqndfnllngls----- 105

OY 126 RKIGS-----CLISANGSEFYGLRPREESNHFLHSCYSVHWLSGVPSGLVIEIGTANK 179

Dd 106 evygnkceevcyymgygsfgrirfpnshlyhsaywhwtqtqpyllstrqialnk 165

OY 180 GSTSYSKCRRPVOKAVLDFTKDTTEFLRIHSKELEFSRGMLLCICVKDDEFDPNPL- 238

Dd 166 gkyislakspvvreaylsqfhedtfmflnaragevvpngcmvl--lltqrscdsedm 223

OY 239 -----DLIDMAINDLIVEGLLEBEKLDSFNIRPFPTPAEAEVKCIYEEGSCETLYLETFFK 293

Dd 224 scftwellamaelaevsgllidedckridtnlpsylfaalevkdlverdsfstidhle--- 280

OY 294 AHVDAAFSIDDYEVNRSHEOIKKEYVASLIRSVEYEPILASHFEALIMPLEFRLAKHAHK 353

Dd 281 -----gfldsevmqendkvrgexktkvzafteplrlanqfgpeimdklydk----- 328

OY 354 VLHMGGCYNNLIIS--LAKPKKSDV 378

Dd 329 -----fthlvsvdleaklphtsl 347

RESULT 6
AAG12686
ID AAG12686 standard; Protein; 379 AA.
XX
AC AAG12686;
XX
Dt 17-OCT-2000 (first entry)
XX
De Arabidopsis thaliana protein fragment SEQ ID NO: 11897.
XX
Km Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
Arabidopsis thaliana.
OS
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825
PR 05-MAR-1999; 99US-0123180
PR 09-MAR-1999; 99US-0123548
PR 23-MAR-1999; 99US-0125788
PR 23-MAR-1999; 99US-0126264
PR 23-MAR-1999; 99US-0126784
PR 01-APR-1999; 99US-0126785
PR 01-APR-1999; 99US-0126782
PR 06-APR-1999; 99US-0128234
PR 08-APR-1999; 99US-0128234
PR 10-APR-1999; 99US-0129845
PR 10-APR-1999; 99US-0130077
PR 21-APR-1999; 99US-0130449
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143242.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145109.
PR 22-JUL-1999; 99US-0145112.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147419.
PR 09-AUG-1999; 99US-0147932.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151062.
PR 27-AUG-1999; 99US-0151068.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155466.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156599.
PR 04-OCT-1999; 99US-0157111.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR	14-OCT-1999	99US-01559684
PR	18-OCT-1999	99US-01595834
PR	21-OCT-1999	99US-01607417
PR	21-OCT-1999	99US-01607417
PR	21-OCT-1999	99US-01607417
PR	21-OCT-1999	99US-01607760
PR	21-OCT-1999	99US-01607760
PR	21-OCT-1999	99US-01608114
PR	21-OCT-1999	99US-01608114
PR	22-OCT-1999	99US-01609880
PR	22-OCT-1999	99US-01609880
PR	22-OCT-1999	99US-01609881
PR	22-OCT-1999	99US-01610494
PR	22-OCT-1999	99US-01610494
PR	22-OCT-1999	99US-01614005
PR	22-OCT-1999	99US-01614005
PR	22-OCT-1999	99US-01614005
PR	22-OCT-1999	99US-01613559
PR	22-OCT-1999	99US-01613559
PR	22-OCT-1999	99US-01613561
PR	22-OCT-1999	99US-01613620
PR	22-OCT-1999	99US-01613620
PR	22-OCT-1999	99US-01613982
PR	22-OCT-1999	99US-01613982
PR	22-OCT-1999	99US-01621442

Query Match	31.0%;	Score 611.5;	DB 21;	Length 379;
Best Local Similarity	38.8%;	Pred. No. 6,3e-53;		
Matches 147;	Conservative	71;	Mismatches 124;	Indels 37;
			Gaps	12

```

Oy      5 EYVHNHEEGTSTANNAS--VYUJLAAKUKPEFLOCRRELYNANPININUKOIVAAOGAS 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      27 kalcmsggdgansyaasnrLqkkvlsmakpvlvryneem--mL-dfrylykvaegss 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      64 GPNYLLIYRDIVOSIRKQCE--EKNELERPPTIOIFLQNDNVSFYKLLSEFYRKLEK 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      85 gqsfslafelintlnvLqbnhks---pelccndlpndentntffkvyfakeml 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      123 ENGRKIGCISLAMPAGFCGRLEFPESMHLNHCYVNHLSQVSGVLYELGANKGSI 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      142 Ltk---sscfvygapsfysrfsrnsnlhhsaylhwlskvekl-----enmkgnl 192
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      183 YSASKCRPPYKALVIOEPRKDTYTLRHSKEFLFSNGRMLLTGICAKUDEDEPNRPD-- 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      193 yltasspsaykaylmgfqkdtfmrlrlrseelivangmvtlflgr--nLlnbrlyudsch 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      240 ----LMDAMINDLYEGELEEEKLEDFNIPPTPSAEEGVCYAEESGCELYLEFP-- 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      252 fctllsmnrltlvgeLveesklakfmmfprdybnvgeLvykqesgefeinelnesngfdl 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      294 AHTDAFASIDDDYVYNSHDLIAEYVASLYRSYVEFLASNGEALNRPDLFHLKNAK 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      312 ghyue-----eddaagne-----agltavsepmllanheIdcltdlkyaivntq 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      354 VLMHGKGINNLTISLAKK 372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      361 hancrktlvaslvsLtkk 379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT	7
AAG37439	
ID	AAG37439 standard; Protein; 379 AA

DT 18-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46035.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN EPI033405-A2.
XX

PD	06-SEP-2000.	
XX		
XX	25-FEB-2000.	2000EP-0301439.
XX		
PR	25-FEB-1999.	9905-0121805.
PR	05-MAR-1999.	9905-0123180.
PR	09-MAR-1999.	9905-0123548.
PR	23-MAR-1999.	9905-0125588.
PR	25-MAR-1999.	9905-0126564.
PR	29-MAR-1999.	9905-0126785.
PR	01-APR-1999.	9905-0127662.
PR	06-APR-1999.	9905-0128234.
PR	08-APR-1999.	9905-0128714.
PR	15-APR-1999.	9905-0129845.
PR	10-APR-1999.	9905-0130077.
PR	21-APR-1999.	9905-0130047.
PR	23-APR-1999.	9905-0130510.
PR	28-APR-1999.	9905-0130891.
PR	30-APR-1999.	9905-0131449.
PR	03-APR-1999.	9905-0132048.
PR	04-MAY-1999.	9905-0132407.
PR	05-MAY-1999.	9905-0132484.
PR	06-MAY-1999.	9905-0132485.
PR	06-MAY-1999.	9905-0132486.
PR	07-MAY-1999.	9905-0132487.
PR	11-MAY-1999.	9905-0132487.
PR	14-MAY-1999.	9905-0134219.
PR	14-MAY-1999.	9905-0134219.
PR	14-MAY-1999.	9905-0134221.
PR	14-MAY-1999.	9905-0134370.
PR	18-MAY-1999.	9905-0134768.
PR	19-MAY-1999.	9905-0134941.
PR	20-MAY-1999.	9905-0135124.
PR	21-MAY-1999.	9905-0135353.
PR	25-MAY-1999.	9905-0135629.
PR	27-MAY-1999.	9905-0136392.
PR	28-MAY-1999.	9905-0136782.
PR	01-JUN-1999.	9905-0137222.
PR	03-JUN-1999.	9905-0137528.
PR	04-JUN-1999.	9905-0137752.
PR	07-JUN-1999.	9905-0137724.
PR	08-JUN-1999.	9905-0138054.
PR	10-JUN-1999.	9905-0138540.
PR	10-JUN-1999.	9905-0138847.
PR	14-JUN-1999.	9905-0139119.
PR	16-JUN-1999.	9905-0139452.
PR	16-JUN-1999.	9905-0139453.
PR	17-JUN-1999.	9905-0139492.
PR	18-JUN-1999.	9905-0139454.
PR	18-JUN-1999.	9905-0139455.
PR	18-JUN-1999.	9905-0139457.
PR	18-JUN-1999.	9905-0139457.
PR	18-JUN-1999.	9905-0139458.
PR	18-JUN-1999.	9905-0139459.
PR	18-JUN-1999.	9905-0139460.
PR	18-JUN-1999.	9905-0139461.
PR	18-JUN-1999.	9905-0139463.
PR	18-JUN-1999.	9905-0139750.
PR	18-JUN-1999.	9905-0139763.
PR	21-JUN-1999.	9905-0139817.
PR	22-JUN-1999.	9905-0139819.
PR	23-JUN-1999.	9905-0140323.
PR	24-JUN-1999.	9905-0140655.
PR	26-JUN-1999.	9905-0140823.
PR	29-JUN-1999.	9905-0140921.
PR	30-JUN-1999.	9905-0141287.
PR	01-JUL-1999.	9905-0141287.
PR	02-JUL-1999.	9905-0142154.
PR	02-JUL-1999.	9905-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 03-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149939.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.0%; Score 611.5; DB 21; Length 379;
Best Local Similarity 38.8%; Pred. No. 6.3e-53;
Matches 147; Conservative 71; Mismatches 124; Indels 37; Gaps 12;

QY 5 EVIHNREGDTSYAKNAS--YNLALAKVPFLEOCIRELTPANLPINCKIRVADIGCAS 63
Db 27 kalcmsggdgsnsrslqkklvismakpvlvrnteeem-mnl-dfpyklvaelgcs 84

QY 64 GNTLLTRDIOVSDKVGGE-EKNELEPRTOIRINDLPONDNSYFKLPSFYKLEK 122
Db 85 gnsflafeflnlnlvlnqhnks--pelaccnolpendntctkvpfnkclmt 141

QY 123 ENGRKISCLISAMPGSFYGRLEPPEBSMHFLHSCYSVHWLSQVPSGLVLELGANKGSI 182
Db 142 ltk--ssctlygagpsfyerltsrnslnhlnssyalnwlsvkpekl-----enknln 192

QY 183 YSSKCCRPVOKAYLDOFTKDTFTFLRIHSKELFSRGRMLLTJCICKVDEDFENPLD--- 239
Db 193 yltsspsaykaylnqfkdftmfllrseelvangrmvltfgr-ntlndplyrdch 251

QY 240 --LIDMAINDLIVEGLLEBEKLDSEFNIFFTPSAEVKCIVEBEGSCILYLF---K 293
Db 252 fctllnsrlrdlveglvseskldatmpfygpnvqelkevlgqesfeneleahgd1 311

QY 294 AYWDAPFIDDDYVVRSHEOIKAEVYASLRSVEPTIASHFGPAIMPDLFRLKNAK 353
Db 312 ghyve---eddfeagrne-----anglrsepmnlahfgeelltdtfdkyuyhvtg 360

QY 354 VHMKGKCYNNLISLAK 372
Db 361 hancnktvslvslck 379

RESULT 8

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0150386.
PR 26-AUG-1999; 99US-0150864.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155559.
PR 29-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156396.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 23-OCT-1999; 99US-0160980.
PR 23-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 30.9%; Score 609.5; DB 21; Length 349;
Best Local Similarity 38.9%; Pred. No. 8, 9e-53;
Matches 146; Conservative 70; Mismatches 122; Indels 37; Gaps 12;

OY 9 MNEGEGDTGYAANKAS-VYLAIAKVPLEBCEIRELLARALPINCICVADICAGSPNF 67
DB 1 ms99gagansysnsrlqkkyismakpviynteeum-mml-dfptyikvaalgses9qns 58
OY 68 LILVRIIVGSIKVGGE-EKNELEPRTIOFLNDLFQDNFVSEKLEFSYRKLEKENGK 126
DB 59 flatefintlnvclqhnkns--pelddcndndpendntftrfypfinkkelmtnk- 114
OY 127 KIGSCLISAMPSSFYGRLEPESNMHLASCVYHMLSOVPSGLVTELGIGANKGSIVSK 186
DB 115 -sacttyagapsfysrlfnsnshlshsyalhwhskvpekl-----enknqnylts 166
OY 187 GCRPPYOKAVLDQFTKDTFTFLRIHSKELFSRGRLMLTICVAVDEDFDPNLD-----L 240

DB 167 sspasaykaylnfqkdfctmfliriseelvsngmvtlfigr-ntlnmopllydcchfcl 225
OY 241 LDMAINDLIVEGLLEEKKIDSNIPFTPSAEVYKCIVEEGSCELLYET--KAHYD 297
DB 226 lsmrlrlvteglvseskldatnmpdybnvqelkevlgkgsfenelesngldghy 285
OY 298 AAFSIDDYPRKSHQIKAEVYASLIRSVEPIIASHGCEAIMPDLEFRLAKNAKVLAM 357
DB 286 e-----eddfaeagne-----anglravsepmlianhgeelldtlfdkyayhvtqhanc 334
OY 358 GKGCYNNLIISLAKK 372
DB 335 rnkttvaalvsllck 349
RESULT 9
AAC37440
ID AAC37440 standard; Protein; 349 AA.
XX
AC AAC37440;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46036.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125786.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0128462.
PR 06-APR-1999; 99US-0128434.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140993.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141844.
PR 02-JUL-1999; 99US-0141854.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143564.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145199.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154029.
PR 16-SEP-1999; 99US-0154179.
PR 20-SEP-1999; 99US-0154179.
PR 22-SEP-1999; 99US-0154139.
PR 23-SEP-1999; 99US-0154586.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query MatchConsistency 30.9%; Score 609.5; DB 21; Length 349;
Best Local Similarity 38.9%; Pred. No. 8.9e-53;
Matches 146; Conservative 70; Mismatches 122; Indels 37; Gaps 12;

9 NMESGSDTSYKKNAS-YNLAKVKPPELOCIIRBLRANPINNCICRVADGASGPN 67
1 mgsdgansysaansrirkkvlsmakprlvrrnteeamm-mtl-dipykylvaalgssgqs 58
68 LITVNDIYQSIDKKGQE-EKNELEPRTQITLNDLFONDNYSYFKLPSFRKLEKNG 126
59 flatelelncrlncqmvnns---peldccclndpendntctfkvpfnkeltmk- 114
127 KIGCSLLIAMGSGTYGRLPPEESMHFLHSCYSVHMVSQVPSGLVIELIGANKSGIYSSK 186
115 ---scctfygagsgstysrlstsrnlhlnssysalhwskvpekl-----enknlylts 166
187 GCRPVOKAYLDQFTKDFTFPLRIHSEKELFSRGMMLLTICIKYDEPDEPNLD-----L 240
167 sspgsaayxaylndqkdfmrlftriseelvangrmvltflgr-nlndpilyrdocbfwcl 225
241 LDMANINDIVEGLLEBEKLDSPNIPFTPSAEKVCIVEBEGSCETLYLEFF---KAHYD 297
226 lnsrlrdlvegllyseakldafmfpdygpnvqelkevlgkgsfelnleeshgfdlghy 285
298 AAFSIDDDYPRSHQEQKAKEYVASLIRSVYPTIASHFGCAIMPDLEFRLKHAAYLHM 357
286 e-----edfseagrne-----angravrsepmiahfgeelldtfdkyhyvtqtanc 334
358 GKGCYNMIIISLAKK 372
335 rnkltvsilvaltkk 349

RESULT 10
AAIG12688
ID AAIG12688 standard; protein: 327 AA.
XX
AC AAIG12688;
AC
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 11899;
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 28-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128232.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128815.
PR 19-APR-1999; 99US-0128815.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

[illegible]


```

PR 22-JUL-1999; 99US-0145087
PR 22-JUL-1999; 99US-0145089
PR 22-JUL-1999; 99US-0145192
PR 23-JUL-1999; 99US-0145146
PR 23-JUL-1999; 99US-0145218
PR 23-JUL-1999; 99US-0145224
PR 26-JUL-1999; 99US-0145276
PR 27-JUL-1999; 99US-0145918
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146389
PR 03-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 06-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148684
PR 16-AUG-1999; 99US-0149175
PR 17-AUG-1999; 99US-0149168
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149902
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066
PR 30-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438
PR 01-SEP-1999; 99US-0151930
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0158293
PR 13-OCT-1999; 99US-0158293
PR 13-OCT-1999; 99US-0158295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768

```

```

PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 22-OCT-1999; 99US-0161040
PR 22-OCT-1999; 99US-0161404
PR 22-OCT-1999; 99US-0161405
PR 23-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161992
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

Query Match 29.7%; Score 586; DB 21; Length 327;
Best Local Similarity 39.6%; Pred. No. 1,9e-50;
Matches 139; Conservative 65; Mismatches 111; Indels 36; Gaps 11;

OY 32 KPFLKCEIRELLRANLPINKIKIVADLCASGPTLTVVDIYOSIDKXGGE-ENKELE 90
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3 kpylvnteeem-mml-dfptylvaeigcsqgnstlaifeitlnctlnvlgqvnms-- 58

OY 91 RPTIQFLNDLFQNDPNSVEKLLPSEYRKLENGKRIKISCLISAMPSSYGLPPESSM 150
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 59 -peidcdclndipendititfitypfinkelmtnk--sctfvgapgsfysrlsfnsl 114

OY 151 HFLHSCYSVHMLSQVPSGLVLELGIGANKGSIYSSKGRPPVOKAYLDQFTKDTFLRI 210
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 115 hlhsayalhwskvpekl-----enngnlyltsspgsaaaykaylndqtkdftmfl 168

OY 211 HSKELFSRGRMLTFCICVDFDEPNPLD-----LDMATNDIVGGLKEEKLDNFNI 264
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 169 rseelavngmvlftfgr-nlndplyrdechfwllnsrlrdlvtvseakldafnm 227

OY 265 PFTPSAEVKEKIVVEEGSCCELLYLETF--KAHYDAFSDIDDYPRSHQIKAEVVS 321
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 228 pfidpvtgvlkeivdkegsfelnleehgfdlghyee-----eddfegsrne-----an 276

OY 322 LIRSVYEPILASHREGAIMPDLFRLAKNAKAYLHMGKCYNNILISLAK 372
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 277 gtlvasepmlshgqeilditlckayavhtqanernktvslvsaltrk 327

RESULT 11
AAG37441
ID AAG37441 standard; Protein: 327 AA.
XX
AC AAG37441;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46037.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

```

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134418.
PR 14-MAY-1999; 99US-0134419.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134770.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139319.
PR 16-JUN-1999; 99US-0139432.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143277.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145098.
PR 21-JUL-1999; 99US-0145098.
PR 22-JUL-1999; 99US-0145098.
PR 22-JUL-1999; 99US-0145098.
PR 22-JUL-1999; 99US-0145098.
PR 22-JUL-1999; 99US-0145112.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147463.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155132.
PR 22-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156639.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0158283
PR 13-OCT-1999; 99US-0158284
PR 13-OCT-1999; 99US-0158285
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 25-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405
PR 25-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161992
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

Query Match 29.7%; Score 586; DB 21; Length 327;
Best Local Similarity 39.6%; Pred. No. 1.9e-50;

Matches 139; Conservative 65; Mismatches 111; Indels 36; Gaps 11;

OY 32 KPFEQCIIRLRLANLPNNKCIKADGASGPNLTVRDIVOSIDKVOE-EKNELE 90
DB 3 kpylvntteemm-mml-dfpylkvaelsgssgnsfllfelfltlavlqvhknks-- 58
OY 91 RPTQIFLNDLONPNFNSFKLEKXGKRIKSGISLAMPSPFGRLPPEESM 150
DB 59 -peldcdlnpndnhtcfkfvfnkeltmk---scsfvypapsfysrlfsmnl 114
OY 151 HFLHSCYSVHMLSOVPSGLVELGIGANKSIVSKGCRPVOKAYLDQFTDFTTFLRI 210
DB 115 hlmsyvalhwlskypckl-----enkgnllytsspspsaykaylqfkdftmfir 168
OY 211 HSKELFSRGMILTICIKVDFDEPNPLD-----LIDMAINDLIVEGILEEKIDSFNI 264
DB 169 rseelvsngmwtlffigr-nlndplyrdchfcllsmelldlvfegivesekldafnm 227
OY 265 PEFMPSEAEVKCIVERESECELYLEFP---KAHYDAASIDDDYVPSHEDIKAEVYAS 321
DB 228 pfygpnvgeikveiksgsfelnelesngfldghye---eddfgaagne-----an 276
OY 322 LIRSVYEPILASHFGALMPDLFRLAKHAKAYVLMKGCICNNLIISLAKK 372
DB 277 glravsepmlahfgeelldtclfdkayhvtqanrcnktvtvslvslckk 327

RESULT 12
ID AAG06670 standard; Protein: 374 AA.
XX AAG06670;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3527.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

OS Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123180.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 11-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135523.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0135629.
XX 27-MAY-1999; 99US-0135621.
XX 28-MAY-1999; 99US-0135622.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140623.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.

PR 18-AUG-1999; 99US-0149426-
PR 20-AUG-1999; 99US-0149722-
PR 20-AUG-1999; 99US-0149723-
PR 20-AUG-1999; 99US-0149929-
PR 23-AUG-1999; 99US-0149902-
PR 23-AUG-1999; 99US-0149930-
PR 25-AUG-1999; 99US-0150566-
PR 26-AUG-1999; 99US-0150884-
PR 27-AUG-1999; 99US-0151065-
PR 27-AUG-1999; 99US-0151066-
PR 27-AUG-1999; 99US-0151080-
PR 30-AUG-1999; 99US-0151303-
PR 31-AUG-1999; 99US-0151438-
PR 01-SEP-1999; 99US-0151930-
PR 07-SEP-1999; 99US-0152363-
PR 10-SEP-1999; 99US-0153070-
PR 13-SEP-1999; 99US-0153758-
PR 15-SEP-1999; 99US-0154018-
PR 16-SEP-1999; 99US-0154039-
PR 20-SEP-1999; 99US-0154779-
PR 22-SEP-1999; 99US-0155139-
PR 23-SEP-1999; 99US-0155486-
PR 24-SEP-1999; 99US-0155659-
PR 28-SEP-1999; 99US-0156458-
PR 29-SEP-1999; 99US-0156586-
PR 04-OCT-1999; 99US-0157117-
PR 05-OCT-1999; 99US-0157753-
PR 06-OCT-1999; 99US-0157865-
PR 07-OCT-1999; 99US-0158029-
PR 08-OCT-1999; 99US-0158232-
PR 12-OCT-1999; 99US-0158369-
PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 13-OCT-1999; 99US-0159295-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159637-
PR 14-OCT-1999; 99US-0159638-
PR 18-OCT-1999; 99US-0159584-
PR 21-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 21-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 22-OCT-1999; 99US-0160989-
PR 23-OCT-1999; 99US-0161404-
PR 23-OCT-1999; 99US-0161405-
PR 25-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 26-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161922-
PR 28-OCT-1999; 99US-0161993-
PR 29-OCT-1999; 99US-0162142-
Query Match 21.9%; Score 432; DB 21; Length 374;
Best Local Similarity 32.6%; Pred No. 8e-35;
Matches 123; Conservative 55; Mismatches 159; Indels 40; Gaps 11;

QY 1 MLEQVLEHMGEGDTYAKNASTYALAKVPEFEOCIRLEFLNANPIN-----KC 53
DB 1 mkterlismkqgkqgdysamnsqagmhar-----smlllleetlelvhlmsasppp 53
QY 54 IKVADIGCAGGNLTLLRVIOVSDKVGQEKNELEKPTIOIFLNDLFQNDENVFKLL 113
DB 54 ftevdigcsggnltvlhldivkhik--rfdaagldppeftafsfclpndfnllfql 111
QY 114 PSFYRLKEK-----NGRKIGSCLISAMPGSFYGRLPFESMHLHSCYSVHWLSQVPSG 168

DB 112 pvlsvntcmeeclaadgnr---syfvagvpsfyrllfpartldffhsafshwlsqvpes 169
QY 169 LVELELGANKSGSIYSKGRCPVQKXVLDQFKDPTFLRHSKELEPSRGRLMTCI-- 226
DB 170 vltirraaynrgvrf-Ingasektctaykrqfgadlaeellrraaevrggamlvclgr 228
QY 227 CKYDERDEPNPLDLDM-----AINDLVEGLEEELKDSFNFPFTPSAEVACIYEBC 282
DB 229 tsydpdcdggagllfcthtfgdwdvlvreglvaaekrdgfnlpyvapslqofkevvdang 288
QY 283 SCEILYLETFKKAHYDAFSDIDYVRSRQIAEYVASLRSYVEPLTSHRGEALMPD 342
DB 289 sfldklvvykxgspdvlnvpepd---asevgra--faascrsvagylveahlgelnsk 342
QY 343 LFRHL-----AKHAKVYL 355
DB 343 lfavvestatshakvyl 359
RESULT 14
AAG06669 standard; Protein; 386 AA.
ID AAG06669
XX AAG06669;
XX AC AAG06669;
XX DE 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 3526.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridization assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX XX 25-FEB-2000; 2000EP-0301439.
XX XX 25-FEB-1999; 99US-0121825.
XX XX 05-MAR-1999; 99US-0123180.
XX XX 09-MAR-1999; 99US-0123548.
XX XX 23-MAR-1999; 99US-0125788.
XX XX 29-MAR-1999; 99US-0126264.
XX XX 29-MAR-1999; 99US-0126785.
XX XX 01-APR-1999; 99US-0128262.
XX XX 06-APR-1999; 99US-0128264.
XX XX 08-APR-1999; 99US-0128714.
XX XX 16-APR-1999; 99US-0129845.
XX XX 19-APR-1999; 99US-0130077.
XX XX 21-APR-1999; 99US-0130449.
XX XX 23-APR-1999; 99US-0130891.
XX XX 28-APR-1999; 99US-0131449.
XX XX 30-APR-1999; 99US-0132048.
XX XX 30-APR-1999; 99US-0132407.
XX XX 04-MAY-1999; 99US-0132484.
XX XX 05-MAY-1999; 99US-0132485.
XX XX 06-MAY-1999; 99US-0132486.
XX XX 07-MAY-1999; 99US-0132863.
XX XX 11-MAY-1999; 99US-0132865.
XX XX 14-MAY-1999; 99US-0134256.
XX XX 14-MAY-1999; 99US-0134218.
XX XX 14-MAY-1999; 99US-0134219.
XX XX 14-MAY-1999; 99US-0134221.
XX XX 18-MAY-1999; 99US-0134370.
XX XX 18-MAY-1999; 99US-0134768.
XX XX 19-MAY-1999; 99US-0134941.
XX XX 20-MAY-1999; 99US-0135124.
XX XX 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138004.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139897.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144315.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145112.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147449.
PR 09-AUG-1999; 99US-0147932.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148343.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151302.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157866.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158323.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 21.9%; Score 432; DB 21; Length 386;
Best Local Similarity 32.6%; Pred. No. 83e-35;
Matches 123; Conservative 55; Mismatches 159; Indels 40; Gaps 11;
OY 1 MELQGLVLANEGSDTSYAKNKNSTYLAALAKYKPLEOCIEELIRANLPIN-----KC 53
DB 13 mKLErLsmxgkqgdysannsqamhar-----smhllleclenhnssasppp 65
OY 54 IKVADLGASGSPNTLLTVRDIVQSIDKVGQEKNELEPTIQFLNDLPONDNSVFRL 113
DB 66 fctvldgssgntvhlidflvkhsk--ridaagldppeftafadpsndfnlfgll 123
OY 114 pSFYKLEKE-----NGRKISCLISAMPGSFYGLRPPESMHFLSCYSVHMLSOVPSG 168
DB 124 plvntcmeciaadgnr--syfvaavgpfyrrllpartldffhsafslhwlsqpes 181
OY 169 LVIEELGICANGSISYSGCRPPVOKAYLDQFTKDETFPIRHSKELFRGHMLTCT-- 226
DB 182 vdrtsaayngarf-lhgagekttafkrfgadlaeflrataaevkrgamflvcigr 240
OY 227 CVYDEFEDEPNLDLDK---AINDLIVEGLIEBKLDSENFIPFSPAEVKCIYEEEG 282
DB 241 tsvpdcgsggallfghfdaedvireglvaakctgfnlpyapsldqdkvvdang 300
OY 283 SCEITLYLTFKAHTDAFSDIDYPRSHQIKAEYASLIRSYTEPIILASHFGAIMP 342
DB 301 sfalklvvygsgplvneppd---asevgra--fascrcvaylvcahlgeelank 354
OY 343 LFHRL---AKHAKVL 355
DB 355 llsrvesratshakvdl 371
RESULT 15
AAG46652
ID AAG46652 standard; Protein: 386 AA.
AC AAG46652;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SRO ID NO: 58715.
DE
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
OS Arabidopsis thaliana.
XX
XX EP103405-A2.
XX
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 23-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134700.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143562.
PR 14-JUL-1999; 99US-0143654.
PR 14-JUL-1999; 99US-0143655.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144385.
PR 16-JUL-1999; 99US-0144386.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148311.
PR 13-AUG-1999; 99US-0148685.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157253.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158233.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159639.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161407.
PR 23-OCT-1999; 99US-0161408.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.9%; Score 432; DB 21; Length 386;
Best Local Similarity 32.6%; Pred. No. 8.3e-35;
Matches 123; Conservative 55; Mismatches 159; Indels 40; Gaps 11;

QY 1 MELOEVHNEGSGDTSYKASNYMLAKYKPELEGCIRELLRANLPNIN-----KC 53
DB 13 mkerllismkgkgqdsynmsqagamhar-----smhlleelenynhnsaapp 65
QY 54 IKVADLOCASGPTLITVDIYQSIDKVGCEKNELEPPTIOIFLNDLFONDSYFKL 113
DB 66 ftavdlgssagantvhlidflvkhnk--rtidaagloppettaifsdpsndntllqll 123
QY 114 PSFYKLEKE-----NGRKISCLISAMPSTFYGRLPESMHFLHSCYSVHMLSOYPSG 168
DB 124 pplysnmcmeeciaadgnr--syfvaqypgsfyrilipartlidffhsafslhwlsqyves 181
QY 169 LVIRLIGANKSGSIYSSKCRPPYOKAYLDQFTKDFTEFLRIHSKELFSRGMILTCT-- 226
DB 182 vldrrsaaylrgvrf-ihgagektctaykrqfgadlaeflraaeavrxgamflvlygr 240
QY 227 CKVDEFEPPNPDLIDLM-----AINDLIVEGLLEEKLDSENFIPFTPSAEVKCIVEEG 282
DB 241 tsavpdlqggagllfsgthgdawdlivreglvaekxrdgnlpyyapsgldfkevvang 300
QY 283 SCCELLYETFKAHYDAFSDIDDDYVBSHEQIKAEVYASLRVYEPILASHFGEAIMP 342
DB 301 sfaldxlvvykqgsplvnepd-----asevgra--fascsvagvlyeahlgeelsnk 354
QY 343 LFHRL-----AKHAKYVL 355
DB 355 lfsrvesratshakdvll 371

Search completed: July 27, 2002, 05:39:32
Job time: 3588 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2002, 04:40:49 ; Search time 19.54 Seconds
(without alignments)
472.511 Million cell updates/sec

Title: US-09-971-020-1

Perfect score: 1972

Sequence: 1 MRLQVLEHMGEDTSTYAR.....KCCYNLLIISLAKPKESDV 378

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	4.8	707	1 US-08-528-122-18	Sequence 18, Appl
2	94	4.8	707	5 PCT-US95-11720-18	Sequence 18, Appl
3	94	4.8	750	1 US-08-325-553-2	Sequence 2, Appl
4	94	4.8	750	2 US-08-394-152A-2	Sequence 2, Appl
5	94	4.8	750	4 US-09-044-668-2	Sequence 2, Appl
6	91.5	4.6	1854	4 US-09-004-838-108	Sequence 108, App
7	89	4.5	475	4 US-09-009-494-6	Sequence 6, Appl
8	89	4.5	475	4 US-09-010-233-2	Sequence 2, Appl
9	89	4.5	475	4 US-09-010-233-6	Sequence 6, Appl
10	87	4.4	856	2 US-08-779-113-2	Sequence 2, Appl
11	87	4.4	856	2 US-08-583-562B-2	Sequence 2, Appl
12	86.5	4.4	442	1 US-08-363-255-11	Sequence 4, Appl
13	86.5	4.4	442	1 US-08-576-444-43	Sequence 4, Appl
14	86.5	4.4	442	1 US-08-576-444-43	Sequence 4, Appl
15	86.5	4.4	442	1 US-08-576-444-43	Sequence 4, Appl
16	85.5	4.3	445	1 US-08-363-255-12	Sequence 5, Appl
17	85.5	4.3	445	1 US-08-363-255-12	Sequence 5, Appl
18	85	4.3	2161	3 US-09-085-199B-9	Sequence 12, Appl
19	84.5	4.3	1140	3 US-09-081-320-3	Sequence 9, Appl
20	84.5	4.3	1140	3 US-09-081-320-3	Sequence 9, Appl
21	84	4.3	3218	4 US-09-564-805-237	Sequence 21, Appl
22	83.5	4.2	636	4 US-09-564-805-237	Sequence 21, Appl
23	83.5	4.2	166	1 US-08-362-453-16	Sequence 16, Appl
24	83.5	4.2	560	3 US-09-176-657-7	Sequence 16, Appl
25	83	4.2	900	1 US-08-425-061-20	Sequence 20, Appl
26	83	4.2	914	1 US-08-825-886-21	Sequence 20, Appl
27	83	4.2	914	2 US-08-825-886-21	Sequence 21, Appl

28	83	4.2	1202	1 US-08-425-061-22	Sequence 22, Appl
29	83	4.2	1202	2 US-08-825-886-22	Sequence 22, Appl
30	83	4.2	1363	1 US-08-425-061-23	Sequence 23, Appl
31	83	4.2	1363	2 US-08-825-886-23	Sequence 23, Appl
32	83	4.2	1852	1 US-08-425-061-24	Sequence 24, Appl
33	83	4.2	1852	2 US-08-825-886-24	Sequence 24, Appl
34	83	4.2	1863	1 US-08-425-061-16	Sequence 16, Appl
35	83	4.2	1863	1 US-08-480-784-2	Sequence 2, Appl
36	83	4.2	1863	1 US-08-483-553-2	Sequence 2, Appl
37	83	4.2	1863	1 US-08-483-553-2	Sequence 2, Appl
38	83	4.2	1863	1 US-08-483-553-2	Sequence 2, Appl
39	83	4.2	1863	1 US-08-488-011B-2	Sequence 2, Appl
40	83	4.2	1863	1 US-08-825-886-16	Sequence 16, Appl
41	83	4.2	1863	2 US-08-603-753D-2	Sequence 2, Appl
42	83	4.2	1863	4 US-09-099-753-2	Sequence 2, Appl
43	83	4.2	1863	4 US-08-850-727-2	Sequence 2, Appl
44	83	4.2	1863	4 US-08-986-106-2	Sequence 2, Appl
45	83	4.2	1863	4 US-09-007-678B-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-08-528-122-18
Sequence 18, Application US/08528122
Patent No. 5726044
GENERAL INFORMATION:
APPLICANT: LO, KIN-MING
APPLICANT: SUDO, YUKIO
TITLE OF INVENTION: GILLES, STEPHEN D.
TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIRIEBOLT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,122
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: FIP-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7000
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 707 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..707
OTHER INFORMATION: /note= "EXTRACELLULAR DOMAIN OF
OTHER INFORMATION: PSMA"
US-08-528-122-18

Query Match 4.8%; Score 94; DB 1; Length 707;
 Best Local Similarity 24.8%; Pred. No. 0.24;
 Matches 53; Conservative 24; Mismatches 61; Indels 76; Gaps 13;

154 HSCYSVHML-----SOVPSGLVIELGIGANK-----G 180
 510 HSYETVELVEKFDPMFKYHLVAQVRGAVFEL---ANSIVLPDCRDYAVVLRKYAD 566
 181 SYSSKGRPPVOKAYIDQFTKFT---TELRIHSEKELFSRGRMLTLCICKYDEDEPNP 237
 567 KITSISMKNHPOEMKTYSVSFDLSFAVKNFTEIASK--FSE-----RILODFKSNP 615
 238 LDLDMAINDLIVELGLEEKLDSEFNIP-----FTTSAEVCICVIEEGSCCELLYL 289
 616 IYLRNM--NDOLM--FLERAFIDPLGLDPRPFYRHVYAPSSHN---KYAG----- 659

290 ETEKAHYDAAFSIDDY-PVRSHEQIKAE-YVAS 321
 660 ESFPGIYDALFIESKVDPSKAMGEVKROIYVAA 693

RESULT 2
 PCT-US95-11720-18
 Sequence 18, Application PC/TUS9511720
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF
 TITLE OF INVENTION: PROTEINS AS IMMUNOPUSINS
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HORWITZ &
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/11720
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: RITCHER, EDWARD R.
 REGISTRATION NUMBER: 27,829
 REFERENCE/DOCKET NUMBER: FIP-001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-248-7100
 TELEFAX: 617-248-7100
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 707 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..707
 OTHER INFORMATION: /note="EXTRACELLULAR DOMAIN OF
 OTHER INFORMATION: PSMA"
 PCT-US95-11720-18

Query Match 4.8%; Score 94; DB 5; Length 707;
 Best Local Similarity 24.8%; Pred. No. 0.24;
 Matches 53; Conservative 24; Mismatches 61; Indels 76; Gaps 13;
 154 HSCYSVHML-----SOVPSGLVIELGIGANK-----G 180

510 HSYETVELVEKFDPMFKYHLVAQVRGAVFEL---ANSIVLPDCRDYAVVLRKYAD 566
 181 SYSSKGRPPVOKAYIDQFTKFT---TELRIHSEKELFSRGRMLTLCICKYDEDEPNP 237
 567 KITSISMKNHPOEMKTYSVSFDLSFAVKNFTEIASK--FSE-----RILODFKSNP 615
 238 LDLDMAINDLIVELGLEEKLDSEFNIP-----FTTSAEVCICVIEEGSCCELLYL 289
 616 IYLRNM--NDOLM--FLERAFIDPLGLDPRPFYRHVYAPSSHN---KYAG----- 659

290 ETEKAHYDAAFSIDDY-PVRSHEQIKAE-YVAS 321
 660 ESFPGIYDALFIESKVDPSKAMGEVKROIYVAA 693

RESULT 3
 US-08-325-553-2
 Sequence 2, Application US/0832553
 Patent No. 5538866
 GENERAL INFORMATION:
 APPLICANT: Ron S.
 APPLICANT: Heston, Warren D.W.
 TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10112
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/325,553
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/973,337A
 FILING DATE: 05 NOV 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 1747/41426
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 664-0525
 TELEFAX: (212) 977-9550
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 750 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-325-553-2

Query Match 4.8%; Score 94; DB 1; Length 750;
 Best Local Similarity 24.8%; Pred. No. 0.26;
 Matches 53; Conservative 24; Mismatches 61; Indels 76; Gaps 13;

154 HSCYSVHML-----SOVPSGLVIELGIGANK-----G 180
 510 HSYETVELVEKFDPMFKYHLVAQVRGAVFEL---ANSIVLPDCRDYAVVLRKYAD 609
 181 SYSSKGRPPVOKAYIDQFTKFT---TELRIHSEKELFSRGRMLTLCICKYDEDEPNP 237
 567 KITSISMKNHPOEMKTYSVSFDLSFAVKNFTEIASK--FSE-----RILODFKSNP 658

DB 238 LDLDMAINDLIVEGLLEKEKIDSFNIP-----FFPSAEVCKIVEEGSCILLYL 289
DB 659 IYLRMM--NQOLM--FLERAFIDPLGLDRFYRHVYVYASSSNN-----KYAG-----702
QY 290 EYFKAHYDAFASIDDDY-PVRSHEQIKAE-YVAS 321
DB 703 ESFPGIYDALFDIESKVDPSKAMGEVYKROIYVAA 736

RESULT 4
US-08-394-152A-2
Sequence 2, Application US/08394152A
Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Isreali, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,152A
FILING DATE: 24-FEB-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 41426-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-394-152A-2

Query Match 4.8%; Score 94; DB 2; Length 750;
Best Local Similarity 24.8%; Pred. No. 0.26;
Matches 53; Conservative 24; Mismatches 61; Indels 76; Gaps 13;

QY 154 HSCYSVHML-----SQVPSGLVLEIGIGANK-----G 180
DB 553 HSYVYTYLVEKEFYDPMKYHLVAOVRGGMFEL---ANSIYLPFCRDYAVVLYKRYAD 609
QY 181 SIYSSKGRPRPVQKAYLDQFTKDF---TFLRIHSKELFSRGRMLLTCLCKYDEDEPNP 237
DB 610 KIYISIMKHPOEMKTYSVSFDLSFAVKNFTETASK--FSE-----RLDDFKSNP 658
QY 238 LDLDMAINDLIVEGLLEKEKIDSFNIP-----FFPSAEVCKIVEEGSCILLYL 289
DB 659 IYLRMM--NQOLM--FLERAFIDPLGLDRFYRHVYVYASSSNN-----KYAG-----702
QY 290 EYFKAHYDAFASIDDDY-PVRSHEQIKAE-YVAS 321
DB 703 ESFPGIYDALFDIESKVDPSKAMGEVYKROIYVAA 736

RESULT 5
US-09-044-668-2
Sequence 2, Application US/09044668
Patent No. 6150508
GENERAL INFORMATION:
APPLICANT: Murphy, Gerald P.
APPLICANT: Holmes, Alton L.
APPLICANT: Holmes, Eric H.
APPLICANT: Tino, William Thomas
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC
TITLE OF INVENTION: FOR THE EXTRACELLULAR DOMAIN OF PROSTATE-SPECIFIC
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/044,668
FILING DATE: 18-MAR-1998
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 8511-0013-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6150508e
US-09-044-668-2

Query Match 4.8%; Score 94; DB 4; Length 750;
Best Local Similarity 24.8%; Pred. No. 0.26;
Matches 53; Conservative 24; Mismatches 61; Indels 76; Gaps 13;

QY 154 HSCYSVHML-----SQVPSGLVLEIGIGANK-----G 180
DB 553 HSYVYTYLVEKEFYDPMKYHLVAOVRGGMFEL---ANSIYLPFCRDYAVVLYKRYAD 609
QY 181 SIYSSKGRPRPVQKAYLDQFTKDF---TFLRIHSKELFSRGRMLLTCLCKYDEDEPNP 237
DB 610 KIYISIMKHPOEMKTYSVSFDLSFAVKNFTETASK--FSE-----RLDDFKSNP 658
QY 238 LDLDMAINDLIVEGLLEKEKIDSFNIP-----FFPSAEVCKIVEEGSCILLYL 289
DB 659 IYLRMM--NQOLM--FLERAFIDPLGLDRFYRHVYVYASSSNN-----KYAG-----702
QY 290 EYFKAHYDAFASIDDDY-PVRSHEQIKAE-YVAS 321
DB 703 ESFPGIYDALFDIESKVDPSKAMGEVYKROIYVAA 736

US-09-004-838-108
Sequence 108, Application US/09004838
Patent No. 6350933
GENERAL INFORMATION:
APPLICANT: Michelmore, Richard W.
APPLICANT: Shen, Kathy
APPLICANT: Meyers, Blake
TITLE OF INVENTION: Procedures and Materials for
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk.
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,838
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Elmhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-07881005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0200
INFORMATION FOR SEQ. ID NO.: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1854 amino acids
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY:
LOCATION: 1..1854
OTHER INFORMATION: /note="RG2J deduced sequence"
US-09-004-838-108

Query Match 4.6% Score 91.5; DB 4; Length 1854;
Best Local Similarity 22.0%; Pred. No. 2, 2; Mismatches 155; Indels 105; Gaps 27;
Matches 92; Conservative 67;

QY 3 LOEYLMNNEGDDTSYKNNASYMLAKVPRFECIRRELLANPNIN-KCIXVADLG 60
DB 820 MEELIRHGSSEGT-----TFPKLLKLLHGLPMLGLCL-VNNAIEPKLVOMK 869
QY 61 CASGPN-TLLIVYDIOVSIDRVGOEERNELEPTIOIFLNDLFONDSVFKLLPSFYRK 119
DB 870 LYSIPGFTSYPRNKLKLSL-----KEEVYIKRDI-----LEIHDMLNLEIWPSTLSR 921
QY 120 LKENGKRI-GSC-LISAMPGSFYRLFPESMHFL-----SCYSVHMLQSVPS 167
DB 922 GEKVKRKIKVNRCKLVN-----LPPHNPMSLHLEBELIVEKSGSIEELFRIDL 972
QY 168 GLVETELGIGANKS-----IYSSKGC--RP-----PVOKAYL---DOFL 201
DB 973 DCASVIGEEDNNSSLRNINVENSMKLRVWRKIGANSRPLRGROVYKRIITIRCKRFT 1032
QY 202 KDFETFRIRHSEKLSRGRML-LTICK-VDEPDENRPLDLDMAINDLYESGLEEKI 259
DB 1033 NTFPI-----ITNFDGLALDELISVDGRNDESDQNO-----EQDQITILEKRTIDERT 1083

QY 260 USF-NIPFTPSAEVKKIVEEGSCCELLYLETFRKAHYDAFSDIDDPVRSHQIAEY 318
DB 1084 DSIENVF--PS-----CLMHSFNLKLLILNRVG-VEVVEIESESPT-SRELVTTH 1134
QY 319 VASLIRSYEPILASHFGGEAIPDLFHLAKAKVLMGKGCYN-NLISLAKPKES 376
DB 1135 -----NOQOPV-----IPNLQHLDLGMDNMIRVK-CSNNKFFTLPKQSSS 1178

RESULT 7
US-09-009-494-6
Sequence 6, Application US/09009494
Patent No. 6150340
GENERAL INFORMATION:
APPLICANT: Black, Michael T.
APPLICANT: Lawlor, Elizabeth J.
APPLICANT: Lewis, Ceti J.
TITLE OF INVENTION: No. 6150340el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,494
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,857
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 60/044,365
FILING DATE: 28-APR-1997
APPLICATION NUMBER: 60/044,366
FILING DATE: 28-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T.
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM50028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
INFORMATION FOR SEQ. ID NO.: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-009-494-6

Query Match 4.5% Score 89; DB 4; Length 475;
Best Local Similarity 20.6%; Pred. No. 0.44;
Matches 80; Conservative 63; Mismatches 160; Indels 86; Gaps 16;

QY 7 LHMNEGDDTSYKNNASYMLAKVPRFECIRRELLA-----NLPRINICIVA 57
DB 121 LHMEDAGKSTHNGEISLVDLNRGTPLIEIVSEPDIRSPKRAYALEKLSIIQYGVGS 180
QY 58 DLGASPNLILVR-DIVOSIDKGOE---KNELERPTIOIFLNDLFONDSVFKLL 113
DB 181 DVKMEG-----SLRDNANISLRPYGQKRGFTKRELNN-----LNSFNYVRKGL 224
QY 114 PSFYKLEKE-NGRKIG-----SCLISAMPGSFYRLFPESM-HFLHSCY 157

Db 225 EYERKOEELNNGEIGOFRRPDESGTILMKRKESDDYRFFPPDVLPIYIDAM 284
Qy 158 SVHMLSOVP-----SGLYIELGIGANKSISYSSKGRPPVOKAYLDOFTKDTFTFLRI 210
Db 285 KERVQOTITPELPERKAKYVNEGL-----PAYDAHVLTTLKREMSDFEES 329
Qy 211 --HSEKLEFSRGMULTICICVDEFEDEPNPLDLDMAINDLIVEGLE-----EKKLD 260
Db 330 TIEHGADVKTLSNMLMG---GVNEYLNKNQVVELDTPKLPENLAGMIKILIEDGTMSKIA 386
Qy 261 SFNIPFTPSAEVFCIVEEGSCIELLETFKRAHYDAFISIDDOYPRSHQIKAEYVA 320
Db 387 KKVPELLAAGNAKOIMEDNGLVOISDEATILKIFYNEALD--NNEQSVEDYKNGKGRAM 445
Qy 321 SLIRSYVEPIASHGEAIMPDLFHRILAK 349
Db 446 FLVQO-----IMKASKGA-NPOLVNOILK 469

RESULT 8

US-09-010-233-2
Sequence 2, Application US/09010233
Patent No. 6200774
GENERAL INFORMATION:
APPLICANT: Black, Michael T.
APPLICANT: Lawlor, Elizabeth J.
TITLE OF INVENTION: No. 6200774el Compounds
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010.233
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,857
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 60/044,365
FILING DATE: 28-APR-1997
APPLICATION NUMBER: 60/044,366
FILING DATE: 28-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T.
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GMS50027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-010-233-2

Query Match 4.5%; Score 89; DB 4; Length 475;
Best Local Similarity 20.6%; Pseq No. 0.44; Indels 86; Gaps 16;
Matches 80; Conservative 63; Mismatches 160;

Qy 7 LHMNEGSDYNSAKNAYMALAKYVPELQICIRRELA-----NLPINKICIXVA 57
Db 121 LHMEDGKSTHNGEYSVLDNROGTPLIEYSEPDIRSPKAVAYLEKLASIIQYIGVS 180
Qy 58 DLGASCPNMLTVR-DIVQSIDKGOBE---KNELEPTIOIFLNDLFQDNFYSYKLL 113
Db 191 DVKNEES-----SLRCDANISURPYGOEKRGTRAELEN-----LSEFYVAKGL 224
Qy 114 PSFYKLEKE--NGRRIG-----SCLISAMPGSFYGRULFPRESN--HFLHSCY 157
Db 225 EYERKOEELNNGEIGOFRRPDESGTILMKRKESDDYRFFPPDVLPIYIDAM 284
Qy 158 SVHMLSOVP-----SGLYIELGIGANKSISYSSKGRPPVOKAYLDOFTKDTFTFLRI 210
Db 285 KERVQOTITPELPERKAKYVNEGL-----PAYDAHVLTTLKREMSDFEES 329
Qy 211 --HSEKLEFSRGMULTICICVDEFEDEPNPLDLDMAINDLIVEGLE-----EKKLD 260
Db 330 TIEHGADVKTLSNMLMG---GVNEYLNKNQVVELDTPKLPENLAGMIKILIEDGTMSKIA 386
Qy 261 SFNIPFTPSAEVFCIVEEGSCIELLETFKRAHYDAFISIDDOYPRSHQIKAEYVA 320
Db 387 KKVPELLAAGNAKOIMEDNGLVOISDEATILKIFYNEALD--NNEQSVEDYKNGKGRAM 445
Qy 321 SLIRSYVEPIASHGEAIMPDLFHRILAK 349
Db 446 FLVQO-----IMKASKGA-NPOLVNOILK 469

RESULT 9

US-09-010-232-6
Sequence 6, Application US/09010232
Patent No. 6248557
GENERAL INFORMATION:
APPLICANT: Black, Michael T.
APPLICANT: Lawlor, Elizabeth J.
TITLE OF INVENTION: No. 6248557el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010.232
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/937,857
FILING DATE: 07-FEB-1997
APPLICATION NUMBER: 60/044,366
FILING DATE: 28-APR-1997
APPLICATION NUMBER: 60/044,365
FILING DATE: 28-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T.
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GMS50026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-010-232-6

Query Match 4.54; Score 89; DB 4; Length 475;
Best Local Similarity 20.64; Pred. No. 0.44; Mismatches 160; Indels 86; Gaps 16;
Matches 80; Conservative 63;

QY 7 LHNNEGGTSTAKKASNNALAKVFPLEOCIRELLA-----NLPINCKCIKVA 57
DB 121 LHNNEAGKSTKGEISVDLNKRGTPLEIVSEPDINSPEKAYLEKRSITQITGS 180
QY 58 DLCCASGPTLLTVR-DIVOSIDKVGQEE--KNLEERTIQIIFLNDLFONDENVFKLL 113
DB 181 DVMERG-----SLKCDANISLRPYGQEKFGTKAEIKN-----INSFNYRKGL 224
QY 114 PSFYRLKEE--NGRKIG-----SCLISAMPGSFYGRLPFEESM--HFLHSCY 157
DB 225 EYEEKHQBELLNGEIQETRRFDESTGKTLIMRKSGSDDYRFPEDPIVLYIDDAM 284
QY 158 SVHMLSQV-----SGLVIEIGIGANKSGSIYSSKGRPPVQKAYIDQTKDFTFLRI 210
DB 285 KERVROTIELPDERKAKAYVNEGL-----PAYDAHVLITLTKMSDFEES 329
QY 211 ---HSEKELSRKMLTICICIKVDFEPNLDLMDAINDLIVGLE-----EKKID 260
DB 330 TIEHGADVLTISNMLG---GVNITLKNQVLELDKLPENLGMKRLLEDGTSSKIA 386
QY 261 SFNIPFTPSAEVKCIYEEGSCGCIYLETFAKHAADAFTDDYVPSRHEQIKAEYVA 320
DB 387 KVFPEPLAKGNAQIMEDNGLVQISDENTLTLLKTVNEALD--NNEQSVEDYKNGKKGAMG 445
QY 321 SLIRSYEPILASHGEAIMDEFLRLAK 349
DB 446 FLVGO---IMKASKGA--NPQALNOLLK 469

RESULT 10
US-08-779-113-2
Sequence 2, Application US/08779113
Patent No. 5948891
GENERAL INFORMATION:
APPLICANT: Staunton, Donald E.
APPLICANT: Harris, Edith S.
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
TITLE OF INVENTION: Binding
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,113
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 5948891and
REGISTRATION NUMBER: 39,302
REFERENCE/DOCKET NUMBER: 27866/33773
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-779-113-2

Query Match 4.48; Score 87; DB 2; Length 857;
Best Local Similarity 18.78; Pred. No. 1.9; Mismatches 60; Indels 54; Gaps 4;
Matches 31; Conservative 21;

QY 247 DLVIEGLEBEKLDSENPFTPSAEVKCIYEEGSCGCIYLETFAKHAADAFTDDY 306
DB 110 DLAEIG-----PSKAEISCIDKDKDGTCTVYTLPTLPEDYSILVKYNDKH 153
QY 307 -----PVRSHEQIKAEYVASLIRSYEPILAS----- 333
DB 154 IPGSPYAKITDSSRCQVKGSAADFLDISETDLSLFTASIRAPSGRDEPCLRLP 213
QY 334 --HFGAIMPDLFHLRLAKHAQVLMGKGCYNNLIISLAKKPEKSD 377
DB 214 NNHGISFLP---REVGEHLVSIKKNGNHVAHNSPVSIMVQSEID 256

RESULT 11
US-08-583-562B-2
Sequence 2, Application US/08583562B
Patent No. 5922570
GENERAL INFORMATION:
APPLICANT: Staunton, Donald
APPLICANT: Harris, Edith
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
TITLE OF INVENTION: Binding
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,562B
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 858 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-583-562B-2

Query Match 4.48; Score 87; DB 2; Length 858;
Best Local Similarity 18.78; Pred. No. 1.9; Mismatches 54; Gaps 4;
Matches 31; Conservative 21;

QY 247 DLVIEGLEBEKLDSENPFTPSAEVKCIYEEGSCGCIYLETFAKHAADAFTDDY 306

DB 110 DLAIEG-----PSKAEISCIQKNDGCTCVTLPTLPDGYSLVKYDKH 153
QY 307 -----PVSHEDQIKAEVYASLIRSYVEPIIAS----- 333
DB 154 IPGSPFATKTSDSRCSQVYKGSADPFLDISPTDLSITSLIKASPGRDEPCLLRP 213
QY 334 --HGEKAMPDLFHLAKHAKVLMGKGCYNLLISLAKKPEKD 377
DB 214 NNHIGISFIP---REVGEHLVSIKKNHNVANSFVSLMVQSEIGD 256

RESULT 12
US-08-363-255-4
; Sequence 4, Application US/08363255
; Patent No. 5783386
; GENERAL INFORMATION:
; APPLICANT: JACOBS, JR., WILLIAM R.
; APPLICANT: BLOOM, BARRY R.
; APPLICANT: COLLINS, DESMOND M.
; APPLICANT: de LISLE, GEOFFREY W.
; APPLICANT: PASCOPELLA, LISA
; APPLICANT: KAWAKAMI, RIKU P.
; TITLE OF INVENTION: MYCOBACTERIA VIRULENCE FACTORS AND A
; TITLE OF INVENTION: NOVEL METHOD FOR THEIR IDENTIFICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,255
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20002.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-363-255-4

Query Match 4.4%; Score 86.5; DB 1; Length 442;
Best Local Similarity 23.4%; Pred. No. 0.74; Indels 67; Gaps 12;
Matches 58; Conservative 34; Mismatches 89;
QY 55 KYADLGCAAGPNTLLTVRDVOSIDV-----GOEKNLEPRTIOFLNDLFQNDPNSVF 110
DB 135 QYAAAGATADP-----VKDYLOIGVPLLNADQVELAKRIENGFLFADKLAN-----D 185
QY 111 KLPSFYRKLE--KENGKIGSCLISA-----MPGSFYGRLPFESNHLHSCYSVHW 161
DB 186 KLAPFKRELEIILIEDGRRAKNHLLLEANLRLVSLAKRYTGR-----GMLFL----- 232
QY 162 LSOVPSGLVIELIG-----ANKGSIVSSKGRPPVQKAVLDOFTKDFTFELR 209

DB 233 -----DLIOGNGILIRAVEKEDYTKGKESTYATWMIROAITRAMADOAR--TIRIP 283
QY 210 ISKELFSR-----GRMLTFCICVDEPDEPNPDLDMALINDLIVGLLEEKLSDFNP 265
DB 284 VMVEVINKLAKRVORQMLQ-----DLGREPPEELAKEL--DWTPEKVEVQKYGREIS 336
QY 266 FFPESAEE 273
DB 337 LHTPLGED 344

RESULT 13
US-08-363-255-11
; Sequence 11, Application US/08363255
; Patent No. 5783386
; GENERAL INFORMATION:
; APPLICANT: JACOBS, JR., WILLIAM R.
; APPLICANT: BLOOM, BARRY R.
; APPLICANT: COLLINS, DESMOND M.
; APPLICANT: de LISLE, GEOFFREY W.
; APPLICANT: PASCOPELLA, LISA
; APPLICANT: KAWAKAMI, RIKU P.
; TITLE OF INVENTION: MYCOBACTERIA VIRULENCE FACTORS AND A
; TITLE OF INVENTION: NOVEL METHOD FOR THEIR IDENTIFICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,255
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20002.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-363-255-11

Query Match 4.4%; Score 86.5; DB 1; Length 442;
Best Local Similarity 23.4%; Pred. No. 0.74; Indels 67; Gaps 12;
Matches 58; Conservative 34; Mismatches 89;
QY 55 KYADLGCAAGPNTLLTVRDVOSIDV-----GOEKNLEPRTIOFLNDLFQNDPNSVF 110
DB 135 QYAAAGATADP-----VKDYLOIGVPLLNADQVELAKRIENGFLFADKLAN-----D 185
QY 111 KLPSFYRKLE--KENGKIGSCLISA-----MPGSFYGRLPFESNHLHSCYSVHW 161
DB 186 KLAPFKRELEIILIEDGRRAKNHLLLEANLRLVSLAKRYTGR-----GMLFL----- 232
QY 162 LSOVPSGLVIELIG-----ANKGSIVSSKGRPPVQKAVLDOFTKDFTFELR 209
DB 233 -----DLIOGNGILIRAVEKEDYTKGKESTYATWMIROAITRAMADOAR--TIRIP 283

OY 210 ISKELFSR-----GRMLTCTCKVDEPNPLDLDMAINDLIVEGLLEEKDLSFNIP 265
DB 284 VHVAVYINKLARVOROMLO-----DLGREPTPELAKEL--DMTPKYEIVOKYGEPIIS 336
OY 266 FTSPSADE 273
DB 337 LHTPLGED 344

RESULT 14
US-08-676-444-43
Sequence 43, Application US/08676444A
Patent No. 6294325
GENERAL INFORMATION:
APPLICANT: Wetmur, James G.
TITLE OF INVENTION: CLONING AND EXPRESSION OF THERMOSTABLE
FILE REFERENCE: MSM95-02
CURRENT APPLICATION NUMBER: US/08/676,444A
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FASTSEQ for Windows version 4.0
SEQ ID NO 43
LENGTH: 649
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-08-676-444-43

Query Match 4.4%; Score 86; DB 4; Length 649;
Best Local Similarity 20.4%; Pred. No. 1.6;
Matches 55; Conservative 44; Mismatches 104; Indels 68; Gaps 10;

OY 144 LPEEEMHFLHSCYVHMLSOVPSGLVIEIGANGSIYSKGCGRPVOKAYIDQFTKD 203
DB 393 LFAKETLIDRLTPALHFAERKP-----ANYDQDHELDIASIDKAY-DLRE 441
OY 204 FTT-----FLIRSKELFSRGMLTCT-----CKVDEPNPLDLDMAINDL 248
DB 442 EASSPELEFFGOMGYLFAAGDGIYIIDQNAQERKYEIYE--SIGNDOSQOL 499
OY 249 IVGGLLEEKDLSFNIPFTPSAEVACIYEERG-----C 284
DB 500 LVPIYIEFPADALRLKRMPLLEEVEFLAEYGENQFLRHPIMAAEELISGIYEMC 559
OY 285 EILYL--ETFKAHYDAF-----SIDDDYVNSHEQIKAEVYASLIRSVY----EP 329
DB 560 DMLLRKEVSIKKYAEALAIMSCRSIKANHRIDHSARQLLYOLSCDNPNYCPHGRP 619
OY 330 ILASHGEAIMPLDPLRLAKAFAKVLHMGK 359
DB 620 VLV-HFTKSDMEKMFRIQENHTSLRELK 648

RESULT 15
US-08-363-255-5
Sequence 5, Application US/08363255
Patent No. 5783386
GENERAL INFORMATION:
APPLICANT: JACOBS, Jr., WILLIAM R.
APPLICANT: BLOOM, BARRY R.
APPLICANT: COLLINS, DESMOND M.
APPLICANT: de Lisle, GEOFFREY W.
APPLICANT: PASCOBELLA, LISA
APPLICANT: KAMAKAMI, RIKU P.
TITLE OF INVENTION: MYCOBACTERIA VIRULENCE FACTORS AND A
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOEISTER
STREET: 755 Page Mill Road
CITY: Palo Alto

STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,255
FILING DATE: 23-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, GLADYS H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 25237-20002.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-363-255-5

Query Match 4.3%; Score 85.5; DB 1; Length 445;
Best Local Similarity 23.4%; Pred. No. 0.97;
Matches 58; Conservative 34; Mismatches 89; Indels 67; Gaps 12;

OY 55 KVALGCGASGPNLTIVRDIYOSIDKV---GOEKNELERPTIOIFNDLFONDFNSVF 110
DB 138 QVAVAGATADP-----VKDYTLKQICKVPLNABQEVLEAKRIEAGLFAEDKLAN---AD 188
OY 111 KLPSFYRKLK--KENGRIKSGCLISA-----MPSFYGRFPRESMHLFHCYSVHW 161
DB 189 KLAKLRLRELIITAEDEGRANKHLLLEALRLVYSLAKRYGR-----GMLL----- 235
OY 162 LSOVPSGLVIEIG-----ANKSIISSKGCGRPVOKAYIDQFTKDTFTFLR 209
DB 236 -----DLIOGMLGIRAVEKFDYTKGKFSYATWTWITQATRAMADOAR---TRIP 286
OY 210 ISKELFSR-----GRMLTCTCKVDEPNPLDLDMAINDLIVEGLLEEKDLSFNIP 265
DB 287 VHVAVYINKLARVOROMLO-----DLGREPTPELAKEL--DMTPKYEIVOKYGEPIIS 339
OY 266 FTSPSADE 273
DB 340 LHTPLGED 347

Search completed: July 27, 2002, 05:40:07
JOB time: 3558 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2002, 04:45:09 ; Search time 37.67 Seconds
(without alignments)
964.209 Million cell updates/sec

Title: US-09-971-020-1
Perfect score: 1972
Sequence: 1 MEQEVLMHNEGEGDTSYAK.....KGCYNMILSLAKPKESDV 378

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614.5	31.2	371	2 E85430	hypothetical prote
2	577.5	29.3	359	2 F84513	hypothetical prote
3	533	26.5	323	2 B86329	hypothetical prote
4	439	22.3	351	2 A86285	protein F91.6 [lm
5	423.5	21.5	363	2 E96703	hypothetical prote
6	411.5	20.9	619	2 T05062	hypothetical prote
7	404.5	20.5	379	2 T51784	ATPP-1 like protein
8	401	20.3	348	2 T51783	ATPP-1 like protein
9	400	20.3	348	2 T51781	protein kinase Atpp
10	127.5	6.5	534	2 C90542	hypothetical prote
11	118	6.0	439	2 S51378	probable membrane
12	107.5	5.5	999	2 F72453	probable cytochrom
13	106	5.4	986	2 T38205	probable cytochrom
14	103.5	5.2	946	2 G71617	conserved hypotet
15	102.5	5.2	1436	2 B81704	SRRA anligen/bapal
16	102.5	5.2	708	2 T39742	hypothetical prote
17	102.5	5.2	2210	1 RRAETV	hypothetical prote
18	102	5.2	344	2 H83928	genome polypeptid
19	101	5.1	405	2 T12824	hypothetical prote
20	101	5.1	2505	1 XYRTFA	fatty-acid synthas
21	99.5	5.0	997	2 T15243	hypothetical prote
22	98	5.0	663	2 E96705	hypothetical prote
23	98	5.0	2185	1 GNNYBT	genome polypeptid
24	97.5	4.9	473	2 F69255	hypothetical prote
25	97	4.9	755	2 T41912	structural phospho
26	97	4.9	1294	2 T04278	hypothetical prote
27	97	4.9	1662	2 T01893	hypothetical prote
28	97	4.9	2185	1 GNNYSV	genome polypeptid
29	97	4.9	2185	1 GNNYSH	genome polypeptid

30	96.5	4.9	778	2 A64656	hypothetical prote
31	96.5	4.9	978	2 T00336	hypothetical prote
32	96	4.9	2201	1 GNNYAS	genome polypeptid
33	95.5	4.8	242	2 E97000	lipase-esterase re
34	95.5	4.8	1127	2 S47445	MDM1 protein - yea
35	95	4.8	1407	1 BMYW1	TARA box-binding p
36	95	4.8	2185	1 GNNYB3	genome polypeptid
37	94.5	4.8	401	2 C70159	translation elonga
38	94.5	4.8	865	2 B64618	preprotein translo
39	94.5	4.8	343	2 T42207	breast cancer susc
40	94	4.8	648	2 T41915	hypothetical prote
41	94	4.8	750	2 A56881	prostate-specific
42	93	4.7	1927	2 T25604	hypothetical prote
43	92.5	4.7	260	2 T41554	hypothetical prote
44	92.5	4.7	329	2 T25931	hypothetical prote
45	92.5	4.7	684	2 T32024	hypothetical prote

ALIGNMENTS

RESULT 1
E85430
hypothetical protein At4g36470 (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_Change 16-Feb-2001
C:Accession: E85430
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp.
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID: 20083488
A:Accession: E85430
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STOS>
A:Cross-references: GB:NC_001268; NID:97270595; PIDN:CAB80313.1; GSPDB:GN00140
C:Genetics:
A:Gene: At4g36470
A:Map position: 4

Query Match	Best Local Similarity	Score	DB 2: Length	DB 2: Indels	DB 2: Gaps	DB 2: Matches	DB 2: Mismatches	DB 2: Indels	DB 2: Gaps
139	39.5%	614.5	371	23	8	139	126	23	8
QY	2 ELOEVLMHNEGEGDTSYAKNAS-YNALAKYKPPLEOCIRELIANLPINIKCRIVADLG 60								
DB	5 DMEREFYMTGDKETSYARNSSLOKASDTRAKHTLETLQOLYKTRP--KSLGIDLG 61								
QY	61 CASGPTLLTVRDIVOSIDKVGQEKNELEPITQIFLNDLPONDNSVFKLPSFYKRL 120								
DB	62 CSSGPTLLSTITPFIKTQVVAHNRPIPIPLREFSIFLNDLPONDNSVFKLPSFYKRL 121								
QY	121 EKENGRIKISC--LISAMPGSFYGRLLPPEESMFLSCSYVHMLSOVPSGLVIELGIA 177								
DB	122 KRDNNN--GCPSPVFLAAYRGSFYGRLEPENTIFHYVASHLSHMTSKVPTALYDOCKSI 179								
QY	178 NKGSIYSKCRPPVQKAYLDQFTKPTIRIKSKELFRGPMILTCCK--VDEPDER 235								
DB	180 NKGVCISCSISSEAVAKACISQFKEDSFIFLCRSKEMVSAGMVAIIIGREGPDHVRG 239								
QY	236 NPL--DILDNAINDLIVEGLLEEKLDSEFNIPFTPSAEVKCIYEEBSCBILYLETFK 293								
DB	240 NSFWEFLSISIDLVAAOGETEEELIDSYDHNHFAAPADEIRGEVDKGSFELRLMLE 299								
QY	294 AHYDAASISD-DYPRSHQIKAVYASLIRSYEPIILASHGCAIPDPEF 344								
DB	300 VKDKGNTGEDIYSY-----GRAVAKTVRAVOESMLVOHGEKILDKLF 342								
RESULT 2									
F84513	hypothetical protein At2g14060 (imported) - Arabidopsis thaliana								

C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84513
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Croitor, I.A.; Shen, M.; Vanden, S.E.; Umayam, U.; Tallon, L.;
 Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-766, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A64420; MUID:20083487
 A:Accession: F84513
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-359 <STO>
 A:Cross-references: GB:AE002093; NID:94388826; PIDN:AAD19781.1; GSPDB:GN00139
 A:Gene: At2g14060
 A:Map position: 2

Query Match 29.3%; Score 577.5; DB 2; Length 359;
 Best Local Similarity 37.6%; Pred. No. 1,2e-36;
 Matches 142; Conservative 72; Mismatches 131; Indels 33; Gaps 12;

9 MNGEGDTSTAAKASYNAL-AKYPELEOCIRELLRANLPINNCIVVADIGASGPT 67
 1 MKGTDGHSYAINSHKORSVFEIOLPIVINYEMLLK--GFPCIVADIGCSTQNT 58
 68 LLIYVDIVOST-DKVGQEKNELEPPTIQLFNDLFONDENSVFLLPSEYRKLEKENG 126
 59 VLAMSLAATITESTYQMSKNP---PEIDYLANDPENFTFLFHSFEKLRPE-- 112
 127 KIGSCLISAMPSEYGRPLPEESMHFLHSCYSVHMLSOVPSGLVETLIGANKSITYSK 186
 113 VKRMVSVGPSEFYSRLPEPKSLHFAVSAFSIHMSLRIPDL-----ESNFTSHIKY 166
 187 GCRPVQAKYLDQFTKPTFLRIHSEKLSRGMILTCI-CKVDEF--DEPNPDL 242
 167 PYPSNVKYSYLNQFIDSLFLKMRSEVYVNGHMLFTVGRKVSDTLSKPCQWMSLS 226
 243 MAINDIVGLLEEKLDSEFNIPFTPSAEVACIYEESGCELIYETFAKHYDAFSI 302
 227 DCLDLASGPFNDMSKVSFNNPYNNEEVEEFILKGSFEIKRK--DHVYVKI 284
 303 -----DDDPVSHHQIAEYVAALIRSVYEPILASHFCEALIMPLFRLAKNAKYLHM 357
 285 DREDEDEQSLOLEAGIKH--ASMAHCITEPLVAHGEDALIEVFVKHYVAKYLSV 341
 358 GKGCYNN--LIISLAK 372
 342 SNRRKMTLVIVSLTK 359

RESULT 3
 866329
 A:Title: Hypothetical protein FL141.3 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: B86329
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Malt, R.; Marzella,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: B86329
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <STO>

A:Cross-references: GB:AE005172; NID:9795588; PIDN:AAF8406.1; GSPDB:GN00141
 A:Gene: At2g14060
 A:Map position: 1

Query Match 26.5%; Score 523; DB 2; Length 333;
 Best Local Similarity 31.1%; Pred. No. 1,6e-32;
 Matches 122; Conservative 77; Mismatches 101; Indels 92; Gaps 10;

1 MELOEVLHMEGEGDTSTAAKASYNALAKV-KPELEOCIRELLRANLPINNCIKYVADL 59
 1 MEYKRVLMHMKNGESTYAKNSTAOSNITISLGRVMDALKIMSN--SEISSIGIDL 58
 60 GCASGPTLLTVVDIVQSIDKVGQEKNELEPPT--IOLFLNDLFONDENSVFLLPSEY 117
 59 CCSSGPNLSLISNIVDTIHLCP---DLDRPVLELVSINLDPNDPNYICASLPFY 114
 118 RLKLENGRKIG-----SCLISAMPSEYGRPLPEESMHFLHSCYSVHMLSOVPSG 168
 115 ---DRVNNKNEGIGFGRGEGECFVSAVPSFGRLPFRSLHFAVHSSSLHMLSKITG 171
 169 L-----VIEIGIGANKSITYSSKGCGRPVQAKYLDQFTKPTFLRIHSEKLSRGM 223
 172 SHNRRELSKGTSSSYVHGRC----- 195
 224 TCICVDFEPEPNLIDMAINDLIVGGLLEEKLDSEFNIPFTPSAEVACIYEESGS 283
 196 -----IGTIEEKIDAFNAPYASSELMLVTEKGS 229
 284 CEIYLETFAKHYDAFSIDDDY--VSHHQIKA--EYVASLIRSVYEPILASHFQAI 339
 230 FSIRLESPIDDEGSISEESYDVIKSRPEALASGRVSNTRIVVEPMLEPTFGNV 289
 340 MPDLFRLAKNAKYLHMKGCYNNLIISLAK 371
 290 MDELFEKAKIVGEYFYVSSPRAYIVLSLR 321

RESULT 4
 A86285
 A:Title: Protein F91.6 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A86285
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon,
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Malt, R.; Marzella,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: A86285
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-351 <STO>
 A:Cross-references: GB:AE005172; NID:95103811; PIDN:AAD39641.1; GSPDB:GN00141
 A:Gene: F91.6
 A:Map position: 1

Query Match 22.3%; Score 439; DB 2; Length 351;
 Best Local Similarity 30.8%; Pred. No. 4,8e-26;
 Matches 111; Conservative 63; Mismatches 140; Indels 46; Gaps 9;

9 MNGEGDTSTAAKASYNALAKV-KPELEOCIRELLRANLPIN--KCIKAVADIGASGP 65
 1 MNGDGASSTAAKASYNALAKV-KPELEOCIRELLRANLPIN--KCIKAVADIGASGP 57

A: Molecule type: DNA
 A: Residues: 1-379
 A: Cross-references: EMBL:AL391254
 A: Experimental source: cultivar Columbia; BAC clone F28D10
 C: GenBank accession: F28D10_60
 A: Map position: 3
 A: Introns: 23/3; 235/3
 A: Note: F28D10_60

Query Match 20.5%; Score 404.5; DB 2; Length 379;
 Best Local Similarity 28.5%; Pred. No. 2.3e-23;
 Matches 109; Conservative 77; Mismatches 138; Indels 59; Gaps 13;

QY 12 GGGTSTAKNASTNLAALAVKFFLEDCIRELL--RANLPINCKIKVADLCAGSPVTL 69
 DB 10 GGGNSTRDHSKTOGALVEA--AKEKINEKISTKLDIDFTSLVNIADFCSSGSPVTF 66
 QY 70 TVRDIVOSIDKVGQEKNELEPPTQIFLNDLFONDNSVFKLLP--SFYRKLKENGK 127
 DB 67 AVQTLIDAVENKKYKESN--IE--FOVFNDSSNNDENTLFTLPPARLY----- 112
 QY 128 IGSCLISAMPGSFYGRLPPEESMHFLHSCYSVHLSQVPSGVVIELGIGANKGSIVSKG 187
 DB 113 ---FAGVYGSFSGFGRVLPRLNSLHGVSAYSLHFISIKPEYVDRDSPVWNR--DIHCS-G 166
 QY 188 CRPPVOKAYIDQFTKFTFLRIHSEKELFSRGMLLTCK--VDEPD--EPNPLDLD 242
 DB 167 SKREKAKLYIGQYKIDVGSFLNARQELVSGLLLLGSCRPNQVMEYEGMMIDFIG 226
 QY 243 MAINDLIVEGLLEEKLDSEFNIPFTPSAEVVKCIYEESCEILYLEFNAHYDAFSL 302
 DB 227 ASLMEINOGILIOQKLDLDFKPIYAPQADELKOIIDNOCFTIEVFENI--IHAKGEYPL 285
 QY 303 DDQYVPSHQIKAQVYASLSVYEPILASHFGEALMPDLFHLAKHAKVLA----- 356
 DB 286 DPEFLTVSFKYTVGGSVASL-----FQDGMKTEFLVAKETOELPOLAKAK 333
 QY 357 -----MGKCYNNLTISLAK 371
 DB 334 PGMOYLIVLRNCFHDLFSLKR 356

RESULT 8
 T51783
 Atpp-like protein - Arabidopsis thaliana
 N: Alternate names: protein F28D10_50
 C: Species: Arabidopsis thaliana (mouse-ear cross)
 C: Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C: Accession: T51783
 R: Delaeny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, August 2000
 A: Reference number: Z25454
 A: Accession: T51783
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-348
 A: Cross-references: EMBL:AL391254
 A: Experimental source: cultivar Columbia; BAC clone F28D10
 C: GenBank accession: F28D10_50
 A: Map position: 3
 A: Introns: 23/3; 238/3
 A: Note: F28D10_50

Query Match 20.3%; Score 401; DB 2; Length 348;
 Best Local Similarity 29.6%; Pred. No. 3.8e-23;
 Matches 106; Conservative 76; Mismatches 124; Indels 52; Gaps 13;

QY 67 TLLVARDIVOSIDKVGQEKNELEPPTQIFLNDLFONDNSVFKLLP--SFYRKLKENGK 124
 DB 64 TETAVQTLIDAVENKKYKESN--IEGIEFQVFNDSSNNDENTLFTLPPARLY----- 115
 QY 125 GRIKISCLISAMPGSFYGRLPPEESMHFLHSCYSVHLSQVPSGVVIELGIGANKGSIVSK 184
 DB 116 ---FAGVYGSFSGFGRVLPRLNSLHGVSAYSLHFISIKPEYVDRDSPVWNR--DIHCS-G 167
 QY 185 SKGCRPPVOKAYIDQFTKFTFLRIHSEKELFSRGMLLTCK--VDEPD--EPNPLD 239
 DB 168 S-GSKREKAKLYIGQYKIDVGSFLNARQELVSGLLLLGSCRPNQVMEYEGMMID 226
 QY 240 LIDMAINDLIVEGLLEEKLDSEFNIPFTPSAEVVKCIYEESCEILYLEFNAHYDAFSL 297
 DB 227 FLSGLMEINOGILIOQKLDLDFKPIYAPQADELKOIIDN--KCFTEKPEKISHAK 283
 QY 298 AAFSIDDYVPSHQIKAQVYASLSVYEPILASHFGEALMPDLFHLAKHAKVLA----- 355
 DB 284 GEPYPLD-----PEYLTSAFKYTVGGSVASLFGQDGMKTEFLVAKETOELPOLAKAK 329

RESULT 9
 T51781
 protein kinase Atpp-like protein - Arabidopsis thaliana
 N: Alternate names: protein F28D10_30
 C: Species: Arabidopsis thaliana (mouse-ear cross)
 C: Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C: Accession: T51781
 R: Delaeny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, August 2000
 A: Reference number: Z25454
 A: Accession: T51781
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-348
 A: Cross-references: EMBL:AL391254
 A: Experimental source: cultivar Columbia; BAC clone F28D10
 C: GenBank accession: F28D10_30
 A: Map position: 3
 A: Introns: 23/3; 238/3
 A: Note: F28D10_30

Query Match 20.3%; Score 400; DB 2; Length 348;
 Best Local Similarity 30.8%; Pred. No. 4.6e-23;
 Matches 106; Conservative 66; Mismatches 126; Indels 46; Gaps 12;

QY 12 GGGTSTAKNASTNLAALAVKFFLEDCIRELL--RANLPINCKIKVADLCAGSPVTL 66
 DB 10 GGGNSTRDHSKTOGALVEA--AKEKINEKISTKLDIDFTSLVNIADFCSSGSPVTF 63
 QY 67 TLLVARDIVOSIDKVGQEKNELEPPTQIFLNDLFONDNSVFKLLP--SFYRKLKENGK 126
 DB 64 TETAVQTLIDAVENKKYKESN--IEGIEFQVFNDSSNNDENTLFTLPPARLY----- 113
 QY 127 KIGSCLISAMPGSFYGRLPPEESMHFLHSCYSVHLSQVPSGVVIELGIGANKGSIVSK 186
 DB 114 KY--FAGVYGSFSGFGRVLPRLNSLHGVSAYSLHFISIKPEYVDRDSPVWNR--DIHCS-G 168
 QY 187 SKGCRPPVOKAYIDQFTKFTFLRIHSEKELFSRGMLLTCK--VDEPD--EPNPLD 241
 DB 169 S-GSKREKAKLYIGQYKIDVGSFLNARQELVSGLLLLGSCRPNQVMEYEGMMID 228
 QY 242 LIDMAINDLIVEGLLEEKLDSEFNIPFTPSAEVVKCIYEESCEILYLEFNAHYDAFSL 300
 DB 228 GSSLDLMEINOGILIOQKLDLDFKPIYAPQADELKOIIDN--KCFTEKPEKISHAK 280
 QY 301 SIDDYVPSHQIKAQVYASLSVYEPILASHFGEALMPDLFHLAKHAKVLA----- 344
 DB 281 GEPYPLD-----PEYLTSAFKYTVGGSVASLFGQDGMKTEFLVAKETOELPOLAKAK 318


```

Db 770 FYPMVKGAEVLD-----DLSSISDLESGIAGLAGEMDNFVPRRAGVSLK 824
Qy 277 IVEEESCCILILEFKAHYA-----AFSIDDYPRSHEDIKAEVYASLRSVPEPIA 332
Db 825 FMEVNEPEVADLR-PEANGAIAIGHVLID-----TTLRGLTNYITVQAPLVS 875
Qy 333 SHEGEAIMDLEFRLAKH-----AAKVLHMGKGCYNLILISLAKPER 375
Db 876 GHGGE-----YHELMVYVYLLNEAMNLRPEEYLALASVMAAG--YN--IASIVSSPSK 923

RESULT 13
T38205
RanBP7/Importin-beta/Cseip homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
C:Accession: T38205
R:Gishchuk, K.; McIntosh, J.R.; Devlin, K.; Churcher, C.; Barrell, B.G.; Rajandream, M.
Submitted to the EMBL Data Library, February 1996
A:Reference number: 221778
A:Accession: T38205
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-986 <DEV>
A:Cross-references: EMBL:Z69730; PIDN:CAA93604.1; GSPDB:GN00066; SPDB:SPAC22H10.03c
A:Experimental source: strain 972H-; cosmid c22H10
C:Genetics:
A:Gene: SPDB:SPAC22H10.03c
A:Map position: 1
A:Introns: 36/1; 776/3; 833/2; 873/1; 930/1; 947/3

Query Match 5.4%; Score 106; DB 2; Length 986;
Best Local Similarity 19.6%; Pred. No. 5.5; Mismatches 130; Indels 148; Gaps 19;
Matches 83; Conservative 62;

Db 529 SVGVKVPILSLDPMLEVL-----SQYASKSSDEALVLLVEAISSAVLDDA 575
Qy 85 EKNELEPPTIOIFLN-----DLFQ-----NDFNSYFKL-LPSFYRKL 121
Db 576 KAELELNSVYPLLEFLVATNADPYIGLIEPTFEDIIHANNNESMCETLPELLOVN 635
Qy 122 KEN-----GRIGSCIL-----SAMPGSFYGRLPPEESMHLHSCYSVHMLSQVPSG-- 168
Db 636 QEDPIVWNGATLSCCLIRAGPSPLPNGFYGLP-----FYKIIITHSGDTE 684
Qy 169 -----LVLELGIGANKGSIYSSKGRPPVQAKAYIDQFTKQFTTF-- 207
Db 685 LILDSLEILKGLLEKPTQLELETFISGSGFQITLF-----TLHQLLDKESDACSFLVG 739
Qy 208 -----LRISKELESRGMILTCICKVDPEDEPNPLDLDMAINDLIVEGLLEEKLDSE 762
Db 740 PILLELADHASQMDLQISILSCSKRLAIAEQPRFQSIIVFAKLIVK-----DEL 791
Qy 263 NIFETPSAEVYKCIIVEEG--SCCEL-----YLETER-----AHYDAFISDD 305
Db 792 GMMHFTLSS-----LLNMGCLTAFVLYMTCDFNVSNNKNTSIIICIMRTIYFSDS- 845
Qy 306 YVPRSHEDIKAEVYASLIR-----SYVEPIASHGEAIMDLEFRLA 348
Db 846 -PLDSDVQVKGELISHSRKRTTSQSKLHPEKYSVSVGEKIL-----RLLSSEPVSL 898
Qy 349 KHA 351
Db 899 KDA 901

RESULT 14
T31617
Sera antigen/papain-like proteinase with active Cys PEB0330c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum

```

```

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: G71617
R:Gardner, M.C.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;
Pierita, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A>Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: G71617
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-946 <GAR>
A:Cross-references: GB:AE001387; GB:AE001362; NID:g3845152; PIDN:AACT1855.1; PID:g384
A:Experimental source: clone 307
C:Genetics:
A:Gene: PEB0330c
C:Superfamily: Plasmodium vivax serine-repeat antigen

Query Match 5.2%; Score 103.5; DB 2; Length 946;
Best Local Similarity 20.6%; Pred. No. 8.1; Mismatches 102; Indels 137; Gaps 18;
Matches 76; Conservative 54;

Qy 20 KNASYNLALAKVAPLEQCI-----RELLRANLPNKKICIKVADLCASGPNLT-LTV 71
Db 284 ENCYHCTLLAKKVDNSNCEFNYSKEAKELINKKKEKTEKGEDEDLSEOKLEESI 343
Qy 72 RDIYOSIDKVGQER-----NELEPPTIOIFLNLPQNDNSYFKLPSFYRKL 121
Db 344 DNLISNTYKIKYESKDKERKSHNNKKELVTE-----ELNSVLKTELLNTYKLL 394
Qy 122 KENR-----KIS-----CLISAMPGSFYGRLPPEESMHLHSCYSVHMLSQVPSG-- 158
Db 395 KEVDRSGMDLHEIGNEIDIRNLLRLKAPG-----ESTVINEKLRNPAICFK 445
Qy 159 -VHMSQVPSGLVLELIGANKGSIYSSKGRPPVQAKV---LDQFTKQFTTFEIRHS 212
Db 446 NIEEMLVN-----KGLLSNKKIONLSTNYNTDLESEYVEYRPF-- 489
Qy 213 KELESGRMILTCICKVDPEDEPNPLDLDMAINDLIVEGLLEEKLDSEFNIPFSPS-- 270
Db 490 -----DDMEFKDMNGVIDLSLFD-----NEKKLS--PYFRANKY 522
Qy 271 -----AEVYKCI-----VEEGSCELIYLFPAKHYDAFISDDYPR-----SHE 312
Db 523 CNNEYCDRKKDKTGCSIKIEVEEGNGCLCMIFASKLHET-----TRCARIGYGH 573
Qy 313 QIRAEVYAS 321
Db 574 RSSSLIVAN 582

RESULT 15
B81704
Conserved hypothetical protein TC0424 [imported] - Chlamydia muridarum (strain N199)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: B81704
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: B81704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1436 <TET>
A:Cross-references: GB:AE002309; GB:AE002160; NID:g7190464; PIDN:AAF39280.1; PID:g719
A:Experimental source: strain N199 (MOPn)
C:Genetics:
A:Gene: TC0424

Query Match 5.2%; Score 103.5; DB 2; Length 1436;

```

Best local similarity 18.68; Pred. No. 14; Matches 74; Conservative 79; Mismatches 145; Indels 99; Gaps 17;

```
OY 2 ELDEVLH-WNEGEGTSTAKNASYNLAAVKEFLQCIPELLRANLPNINKCIKYNADLG 60
Db 381 EIDDAHHEICIEDGISPIYIDP-----QVRIVYQCKREELKDLGKTDEELKPCDLS 432
OY 61 CASGPTLLTVRDIQSID-KVGOEKN--ELERPTQIFINDLFO--NDENSVFKLP 114
Db 433 KIQ-RSVCLFATSVVSLLEGKMGVSEKSIKEIETVRELSTLQIGLSGGITPLID 491
OY 115 SFYKLEKENGKRIKISCLISAMPGSFYGRLEFPRESMHFLHS-----CYSV 159
Db 492 NVHKAIRQ--GRALSNELRQSI-----QLHPRFRHRLQAKVEKLOGFIRDPKMGASAV 543
OY 160 HMLSOVP-----SGVIEIGIGANKSIIYSSKGRPPVOKAYLDQFTKDFTFPL 208
Db 544 HLSQETLEOKRQQLDKLTGIGITLSDMETRYSVK---ETKLFHYMEDFEFETEKFL 599
OY 209 RIHS-----KELFSRGRMLTLCICKYDEPDEPNPLDLMAINDLIYEGILLE 256
Db 600 NSHSAVAESCSLDCSYVELKDCDQALNADIGNIEKVMNPADVESAREFKQLISDLGVQ 659
OY 257 EKLDSENIPTFTPSAEVKCIYEEGSCETLIYETKAYDA-----AFST- 302
Db 660 EQLDQUSVP-----ICERVSGRKLLNTLSHPDLQKKELEKRAALEAFTSG 708
OY 303 -DDDYPRSHQIKA---EYVASLIRSV--YEPIL 331
Db 709 EDPEFPVRKEETLDVSSGYDYLSNLLGKINSFESIL 745
```

Search completed: July 27, 2002, 05:41:00
Job time: 331 sec

Sun Jul 28 10:37:50 2002

us-09-971-020-1.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2002, 05:40:09; Search time 19.69 seconds
(without alignments)
743.321 Million cell updates/sec

Title: US-09-971-020-1

Sequence: 1 MEQAEVLHNEGEEDTSYK.....KGCYNLLISLAKPEKSDV 378

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SWISSPROT_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	106	5.4	986	1	YD43_SCHPO
2	102.5	5.2	2210	1	RPO_TACY
3	101	5.1	2505	1	FAS_RAT
4	100.5	5.1	2144	1	BP28_HUMAN
5	98	5.0	2185	1	POLG_CX3N
6	98	5.0	2203	1	POLG_EC09B
7	97.5	4.9	635	1	ETP1_YABAM
8	97	4.9	755	1	P100_HSV7
9	97	4.9	2185	1	POLG_SVDV
10	97	4.9	2185	1	POLG_SVDV
11	96.5	4.9	960	1	OPAL_HUMAN
12	96.5	4.9	960	1	OPAL_MOUSE
13	96	4.9	2183	1	POLG_CX84E
14	96	4.9	2193	1	POLG_EC30B
15	96	4.9	2201	1	POLG_CXA9
16	95.5	4.8	1127	1	MDM1_YEAST
17	95	4.8	1407	1	TFD2_YEAST
18	94.5	4.8	394	1	EFPU_BORBU
19	94.5	4.8	865	1	SECA_HELPY
20	94	4.8	750	1	GPC2_HUMAN
21	94	4.8	2184	1	POLG_EC01F
22	94	4.8	2185	1	POLG_CX3M
23	94	4.8	2191	1	POLG_EC06C
24	93	4.7	2196	1	POLG_EC05N
25	92.5	4.7	777	1	TBR1_HUMAN
26	92	4.7	2182	1	POLG_CX81J
27	92	4.7	2183	1	POLG_CX84J
28	91.5	4.6	964	1	VATA_PYRHO
29	91.5	4.6	3329	1	BRC2_MOUSE
30	91	4.6	677	1	HSEF_KIUTA
31	91	4.6	2185	1	POLG_CX85P
32	91	4.6	2195	1	POLG_EC11G
33	90.5	4.6	270	1	DMA_SERRA

34	90.5	4.6	559	1	THSA_SUITO
35	90.5	4.6	467	1	RR2_HUMAN
36	89.5	4.5	2238	1	RRP1_BUNYW
37	89	4.5	229	1	Y467_BORBU
38	89	4.5	659	1	SP42_SYNY3
39	89	4.5	1125	1	MPD_BORBU
40	89	4.5	1295	1	EXA2_CLOBO
41	88.5	4.5	1311	1	GAK_HUMAN
42	88.5	4.5	3555	1	YAMB_SCHPO
43	88	4.5	612	1	HS76_YEAST
44	88	4.5	647	1	NRP1_MSEPV
45	88	4.5	1092	1	DP2L_METTH

ALIGNMENTS

RESURF 1
ID YD43_SCHPO STANDARD: PRT: 986 AA.
AC Q10297:1096 (Rel. 34, Created)
DR 31-OCT-2000 (Rel. 39, Last sequence update)
DI 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothesized 111.1 kDa protein C22H10.03c in chromosome I.
GN SPK22H10.03c.
OS Schizosaccharomyces pombe (fission yeast).
OC Schizosaccharomycetes;
OC Eukaryota, Fungi, Ascomycota, Schizosaccharomycetes;
OC Schizosaccharomycetales, Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID:4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:972.
RA Devlin K., Churche C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RT 1. SUGARCOLEA LOCATION: Integral membrane protein (Potential).
CC 1. SIMILARITY: SOME, TO YEAST TGL241W/HRC1004.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL:Z69730; CAA93604.1;
DR InterPro: IPR001494; IPR_NF;
DR Hypothetical protein.
KW SEQUENCE 986 AA; 111078 MW; 640226AB24AB704C CRC64;

Query Match Best Local Similarity 19.68; Score 106; DB 1; Length 986;

Matches 83; Conservative 62; Mismatches 130; Indels 148; Gaps 19;

QY	27	ALAKVVF--LDCQRELLRANLPNINKIKVADLCASGPNLLTVRDIVOSIDK	986	84
DB	529	SVAKVAFILSLDPMLEVL-----SOYAKSKSDPALVILVERISSAVKIDCA	575	
OY	85	EKNLEERPITQIFLN-----DLFQ-----NDPNSVFKL-LPSEYRKLE	121	
DB	576	KAAELGNSYIPLENLVATNADPYICGILEDEPDIITHANNESCEITTELLQVNLN	635	
OY	122	KEN-----GKIGSCL-----SAMPGSTYGRLPFESMHLSCSVHMLSOVPSG--	168	
DB	636	QEDPIMNIGATLISCLIRAGSPPLNGFVGLP-----PVYKTIQIHSGDTE	684	
OY	169	-----LVIEIGCANKGSIYSKSGCRPVOKAYLDFTYDFTTF--	207	
DB	685	LIQLSOLIKLLEKDPQLLEIRISGSGFQITL-----TLHQLDKESDSASCTLVG	739	

```

OY 208 -----LRHSKRELSRGRMLTJCICKYDEDEPNLDDAINDLYEGLEBEKIDSF 262
DB 740 PILLELADNASQMDLOSILSCIKRAIADPRFQSIITYFAKLYK-----DSL 791
OY 263 NIFFTPSAEYKACVEEG--SCETL-----YLETF-----ANYDAFSTDD 305
DB 792 GANHFTSS-----LLNEOGLTAEVMTWCNDFYNSNFENSIICIMTKYSPDS- 845
OY 306 YPVRSHEQIAEYVASLIR-----SYVEPIASHGEGAIMPDLFHRLA 348
DB 846 -PLDVSQYKGLISHSNRITRSQSKLHPEEYSYVSVEKIL-----RLSEEFVSL 898
OY 349 KHA 351
DB 899 KDA 901

RESULT 2
RPRO_TACY STANDARD: PRT: 2210 AA.
AC P20430;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE RNA polymerase (EC 2.7.7.48).
GN
OS Taccaribe virus.
OC Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus.
OX NCBI_TaxID-11631;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89243206; PubMed-2718387;
RA Iapalucci S., Lopez R., Rey O., Lopez N., Franze-Fernandez M.T.,
RA Cohen G.N., Lucero M., Ochao A., Zakin M.M.
RT "Taccaribe virus L gene encodes a protein of 2210 amino acid
RT residues."
RL Virology 170:40-47(1989).
CC -I CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J04340; AAA47901.1;
DR PIR: A31468; RRPXYV.
RW RNA-directed RNA polymerase; transferase.
SQ SEQUENCE 2210 AA; 252230 MW; EACF01C32761024F CRC64;

Query Match 5.28; Score 102.5; DB 1; Length 2210;
Best Local Similarity 21.58; Ptd. No. 6.5; Index 145; Gaps 23;
Matches 89; Conservative 59; Mismatches 121;

OY 34 FLEOCIRELLR-----ANLPNICKIVADLACAGPNTLLTVYDIVQSIDRY 81
DB 413 WENCCKILRRSQOQIWSQISVMRYDPDLKLSIA-----QTSSDPIIRY- 5A 462
OY 82 GOEKNELRPTIQFLNDLFONDFNSVFRLPSFRKLKENGKRGSCSLISAMPGEY 141
DB 463 GGNFNECKHKTFHM-MSDAQOV- -AFKILSS-----VSLSLNSMKTSS 506
OY 142 GRLE---PRESNH- -LHSCYVHMLSGVPGYI-----ELGI 175
DB 507 SRLINKEKSYRFGNVRARCQYQGRF- -LTGGLVILFYKRGKRGSCSYTYCEGV 564
OY 176 GAKGSIYSKCRPPQKALDOFTK-----DFTFLIRHSEKLFSGNM 221
DB 565 LYKGSY---CDP- -KRFPLIFSOEYVEMCDMTWLDPSNDLAVISKELR- -LL 616

```

```

OY 222 LITGICVDEFDSP-----NPDLDD-----MAINDLYEGL----- 253
DB 617 LLSLTCAPSRRNOVFLQGRFLMAYSNQPHVDLSKLEKVECMSSSEVIYORLAVDLFO 676
OY 254 -LEBEKIDSNIFFTPSAEYKACVEEGSCETLYLETFKRAYDAFSTDDYVPVRSHE 312
DB 677 CLLEGVDS--DVEF---ARRKYLANSYLCHLITKET-----PDLRLD 716
OY 313 QIKR--EYVASLIR-----SYVEPIASHGEGAIMPDLFHRLAKHAARVLMGKGC 361
DB 717 QIKCFEFTIEPKIDFNCVIVPNSLNGLTFEAGBMMDGLDKRYSTLT---KDC 767

RESULT 3
PAS_RAT STANDARD: PRT: 2505 AA.
AC P12785; 064717; 009167; 009190;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
GN FASN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89240686; PubMed-2717611;
RA Amy C.M., Witkowski A., Nagert J., Williams B., Randhawa Z.,
RA Smith S.;
RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
RT fatty acid synthase."
RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
RA Beck K.F., Schreggmann R., Stathopoulos I., Klein H., Hoch J.,
RA Schweitzer M.;
RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
RT norvegicus."
RL DNA Seq. 2:359-386(1992).
RN [3]
RP SEQUENCE OF 75-2505 FROM N.A.
RX STRAIN-SPRAGUE-DAWLEY; TISSUE-Mammary gland;
RA Schweitzer M., Takabayashi K., Beck K.F., Schreggmann R.;
RT "Rat mammary gland fatty acid synthase: localization of the
RT constituent domains and two functional polyadenylation/termination
RT signals in the cDNA."
RL Nucleic Acids Res. 17:567-586(1989).
RN [4]
RP SEQUENCE OF 2085-2505 FROM N.A.
RX TISSUE-Mammary gland;
RA MEDLINE-88087240; PubMed-2891707;
RA Nagert J., Witkowski A., Mikkelson J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
RT domain of the rat fatty acid synthetase."
RL J. Biol. Chem. 263:1146-1150(1988).
RN [5]
RP SEQUENCE OF 1921-2324 FROM N.A.
RX TISSUE-Mammary gland;
RA MEDLINE-87246646; PubMed-3109907;
RA Witkowski A., Nagert J., Mikkelson J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
RT protein and its flanking domains in the mammalian fatty acid
RT synthetase."
RL Eur. J. Biochem. 165:601-606(1987).
CC -I- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.

```

THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN ACYL CARRIER PROTEIN.

-1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH - a long-chain fatty acid + (N+1) COA + N CO(2) + 2N NADP(+).

-1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] - CoA + acetyl-[acyl-carrier protein].

-1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] - CoA + malonyl-[acyl-carrier protein].

-1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-carrier protein] - 3-oxoacyl-[acyl-carrier protein] + CO(2) + [acyl-carrier protein].

-1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) - 3-oxoacyl-[acyl-carrier protein] + NADPH.

-1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] - 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.

-1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) - trans-2,3-dehydroacyl-[acyl-carrier protein] + NADPH.

-1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O - [acyl-carrier protein] + oleate.

-1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL, M76767; AAA57219.1; .

CC EMBL, X62888; CAA44679.1; .

CC EMBL, X62889; CAA44680.1; .

CC EMBL, X13415; CAA31780.1; .

CC EMBL, X13527; CAA31882.1; .

CC EMBL, J03514; AAA41144.1; .

CC PIR, A30313; XYRFA.

CC HSSP, P02901; IACP.

CC InterPro: IPR001227; Acyltransferase domain.

CC InterPro: IPR002085; Adh_zn_family.

CC InterPro: IPR000794; Ketoacyl-synth.

CC InterPro: IPR003880; Phosphopantattach.

CC InterPro: IPR000051; SAM_bind.

CC InterPro: IPR001031; Thioesterase.

CC Pfam: PF00698; Acyl_transf_1.

CC Pfam: PF00107; adh_zinc_1.

CC Pfam: PF00109; ketoacyl-synth_1.

CC Pfam: PF02801; ketoacyl-synth_1.

CC Pfam: PF00975; Thioesterase_1.

CC Pfam: PF00975; Thioesterase_1.

CC PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.

CC PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.

CC PROSITE: PS50075; ACP_DOMAIN; 1.

CC Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine; Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP; Pyridoxal phosphate.

CC KMW Pyridoxal phosphate.

CC FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.

CC FT DOMAIN 429 817 ACYL AND MALONYL TRANSFERASES.

CC FT DOMAIN 1629 1857 ENOYL REDUCTASE.

CC FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.

CC FT DOMAIN 2118 2174 ACYL CARRIER (ACP).

CC FT DOMAIN 2202 2505 THIOESTERASE.

CC FT ACT SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

CC FT ACT SITE 161 161 MALONYLTRANSFERASE (BY SIMILARITY).

CC FT ACT SITE 165 165 MALONYLTRANSFERASE (BY SIMILARITY).

CC FT BINDING 165 165 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC FT BINDING 165 165 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC FT BINDING 1765 1780 NADP (KR).

CC FT BINDING 2151 2151 PHOSPHOPANTETHEINE (BY SIMILARITY).

CC FT BINDING 2302 2302 THIOESTERASE (BY SIMILARITY).

CC FT ACT SITE 2475 2475 THIOESTERASE (BY SIMILARITY).

CC FT ACT SITE 2475 2475 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).

CC ACT SITE 878 878 S -> P (IN REF. 3).

CC CONFLICT 871 871 S -> P (IN REF. 3).

CC CONFLICT 1967 1968 MV -> IL (IN REF. 5).

FT CONFLICT 2085 2085 C -> P (IN REF. 4).

FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).

FT CONFLICT 2296 2296 Y -> H (IN REF. 1 AND 5).

SO SEQUENCE 2505 AA; 272647 MW; 5810EC13D37F3114 CRC64;

Query Match 5.1%; Score 101; DB 1; Length 2505;

Best local similarity 21.4%; Pred. No. 10;

Matches 96; Conservative 54; Mismatches 132; Indels 166; Gaps 22;

52 KCIKADIGCAS-----GPNLLTV--RDVOSIDKV-----QEENLEERTI 94

DB 1463 KCLLSLSSSTSHPKLPDPESSBLQVLESDLVNMYVRDGMGAFHFDLEDDKEQTA 1522

QY 95 QIFLNDLFQDNFSV-----FKLPSFYRKLENGKRT---GSCLI 133

DB 1523 HAFVYVLTGRGLASIRVSSPLKMKOPSSGQALCTVYVASL---NFRDILANGKLS 1579

QY 134 SAMPG-----SFYGR-----LPPESMHFLHSCYSVHMLSOVPSGLVIEL 173

DB 1580 DAIPKMASRDCMLGMFSGRDCGRVWGLVPAEGL-ATSVLSPDPLMDVPSMWLEE 1638

QY 174 -----GIGANKSITYSKGR-----PPVQ 193

DB 1639 AASVPVYTVATVSLVYVGRHOGETVLHSGSGVGQGAISIALSGRVTYVSAK 1698

QY 194 KAYLD-OETK-DPTFLIRSKELFRGMLLTCTCKYDEDERPNDLLDMAINDLVE 251

DB 1699 KAYLDARPPOLDTSF--ANSRPTSEOVYLLHTGKG-----VDLVLS----- 1741

QY 252 GLLSEKLDSEFNIDFFTPPSAEVKCIVEEGSCET-----LYETKRAH-- 295

DB 1742 -LAEKELQA-----SVRCIAHQGRLELGRKEDLSNNHPLGMAIFLKNVTEFGI 1788

QY 296 -YDAFSDIDDPYRSHQIKATVYASLIRSYEPILASHEGAILMPDLEHRLA--KHAA 352

DB 1789 LLDALFEGAND---SWREVAELKAGRIDGVKPLCTVFPRAQVEDAFRYMAQGRHIG 1844

QY 353 KVL-----HMKGCYNNLIISLAK 371

DB 1845 KVLGVVREPEEPMALPGAOPTLSAISK 1872

RESULT 4

ID BP28 HUMAN STANDARD; PRT: 2144 AA.

AC 09H5B3; G9NW23;

DT 16-OCT-2001 (Rel. 40, Created).

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein BAP28.

GN BAP28.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

ON [1]

RP SEQUENCE FROM N.A. AND VARIANTS S-1694; A-1854; D-1967 AND 6,2017.

RA Bouquellet L., Chumakov I., Barry C., Cohen-Akenline A.;

RT "A novel BAP28 gene and protein."

RP Patient number W00100669, 04-JAN-2001.

RN [2]

RP SEQUENCE OF 1534-2144 FROM N.A.

RA Chobley V.

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 1777-2144 FROM N.A.

RA Isegaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,

RA Wagatsuna M., Hosogi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,

RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,

RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,

AC P52519; 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large structural phosphoprotein homolog (PPI00).
GN U11.
OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Alphaherpesvirinae; Simplexvirus.
OC NCB1_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO THE LARGE STRUCTURAL PHOSPHOPROTEINS OF HSV-6 AND
HCW UL32.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: U43400; AAC5672.1;
DR Matrix protein; Phosphorylation.
SQ SEQUENCE 755 AA; 86580 MW; 4083744CC3P91DA CRC64;

Query Match 4.9%; Score 97; DB 1; Length 755;
Best Local Similarity 19.8%; Pred. No. 4.2;
Matches 73; Conservative 57; Mismatches 130; Indels 108; Gaps 16;

QY 35 LEOCCRELLRANLPNKKCKVADLGCASGPT-----LTVRDI----- 74
DB 414 INQVDTLSKIDLHNSKVIDI-----VSSPKVNVOLPKNKIDYHSTFLEPENENRON 468
QY 75 -VOSIDKGOEKNELER-----PTQIPLNDLFO-----NDEN 107
DB 469 GVOSSDOLSKNSTNDLOKILERIKITKONNEDIFLPESEKREIVHENLQSPDDEH 528
QY 108 SVFKLIPFVFKLEKNGRKISGLSAMPGSFYRLPESMFLHSCS----- 158
DB 529 NEMSLPPODOKSIRKONKANS---STTLNMITNOVNASMEKESASAKKNQLYVD 585
QY 159 VHWLSQVSGVLTGIGANKGSISYK---GCRPVQKAVLDOFKDTTFLRIHSKE 214
DB 586 VKW---TPSSSLDL---SRNDLQKELFESGLSEKVK---LITDFTTISLEERS 634
QY 215 LFSRGRML-----TCICKVDEPDEPNPLDLLMAINDLIVEGLLEEKLDSPNIFPFP 269
DB 635 L---KDLLEPPKRTDVSNTATNDNNLNKLSRKRDPLEFQNSFTEKKQPVRSFFFLP 690
QY 270 SAEVWCIVEEGSCETILETFKAHYDAFS-----IDDDYVRSHEQIKAE 317
DB 691 NAE-----IOFGDSSL---TGKERTONTIFGASKAQENSGDKLDIDENSVCDDDIYVK 742
QY 318 YVASLIRS 325
DB 743 LVSHLTHS 750

RESULT 9
FOLD_SYDVH STANDARD; PRT; 2185 AA.
ID FOLD_SYDVH
AC P16604;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 to VP4; Core proteins
DE P2A to P2C; P3A; Genome-linked protein VPg; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D

DE (EC 2.7.7.48)).
OS Swine vesicular disease virus (strain H/3 '76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus
OX NCB1_TaxID=12076;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89279274; Pubmed=2543767;
RA Inoue T., Suzuki T., Sekiguchi K.;
RL J. Gen. Virol. 70:919-934(1989).
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS:
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: D00435; BA00337.1;
DR PIR: A1331; GNNYSH.
DR HSSP: P21404; IDAM.
DR MEROPS: C03.001; -;
DR MEROPS: C03.020; -;
DR Interpro: IPR000199; Cys-protease-3C.
DR Interpro: IPR003138; Pico_P1A.
DR Interpro: IPR000081; Pico_P2B.
DR Interpro: IPR002527; Pico_P2B.
DR Interpro: IPR000605; RNA_helicase.
DR Interpro: IPR001205; RNA_pol_P3D.
DR Interpro: IPR001676; Rhv.
DR Pfam: PF00548; Cys-protease-3C; 1.
DR Pfam: PF02226; Pico_P1A; 1.
DR Pfam: PF00947; Pico_P2A; 1.
DR Pfam: PF01552; Pico_P2B; 1.
DR Pfam: PF00073; rhv_3; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR Pfam: PD001125; Cys-protease-3C; 1.
DR Pfam: PD001274; Pico_P2B; 1.
DR Pfam: PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.
FT CHAIN 2
FT CHAIN 69
FT CHAIN 70 330
FT CHAIN 331 568
FT CHAIN 569 851
FT CHAIN 852 1001
FT CHAIN 1002 1100
FT CHAIN 1101 1429
FT CHAIN 1430 1518
FT CHAIN 1519 1540
FT CHAIN 1541 1723
FT CHAIN 1724 2185
FT LIFTD 2
FT ACT_SITE 1687 1687
FT ACT_SITE 1701 1701
FT ACT_SITE 2185 2185
SQ SEQUENCE 2185 AA; 243164 MW; 582285DD9482B3B6 CRC64;

Query Match 4.9%; Score 97; DB 1; Length 2185;
Best Local Similarity 19.2%; Pred. No. 18;
Matches 73; Conservative 62; Mismatches 155; Indels 90; Gaps 16;

ID	OPAL_HUMAN	STANDARD:	PRT:	960 AA.
1C	OPAL133			
1D	OPAL133			
1E	16-OCT-2001 (Rel. 40, Created)			
1F	16-OCT-2001 (Rel. 40, Last sequence update)			
1G	01-MAR-2002 (Rel. 41, Last annotation update)			
1H	Dynamn-like 120 kDa protein, mitochondrial precursor (Optic atrophy 1 gene protein).			
1I	OPAL OR KIA00567.			
1J	Homo sapiens (Human)			
1K	Eukaryota, Metazoa			
1L	Mammalia, Euteleia, Primates, Catarrhini, Homiidae, Homo.			
1M	NCBI_Taxid:9606.			
1N	SEQUENCE FROM N.A.			
1O	TISSUE=Brain;			
1P	MEDLINE=98290545; PubMed=9628581.			
1Q	Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.			
1R	Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.			
1S	DNA Res. 5:31-39(1998).			
1T	[2]			
1U	SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND VARIANT OPAL133-300.			
1V	MEDLINE=20473232; PubMed=11017079;			
1W	Delattre C., Lenaers G., Griffon J.-M., Gigarel N., Lorenzo C., Belanger P., Pellicoquin L., Grosgeorge J., Turc-Carel C., Perret E., Astier-Dequeker C., Laessle L., Arnold B., Ducommun B., Kaplan J., Hamel C.P.			
1X	"Nuclear gene OPAL1, encoding a mitochondrial dynamn-related protein, is mutated in dominant optic atrophy."			
1Y	Nat. Genet. 26:207-210(2000).			
1Z	[3]			
2A	TISSUE SPECIFICITY, AND VARIANTS OPAL1 GLN-290 AND ILE-432 DEL.			
2B	MEDLINE=20473232; PubMed=11017080;			
2C	Alexander C., Votruba M., Pesch U.E.A., Thiselton D.L., Maier S., Moore A., Rodriguez M., Kelner U., Leo-Kottler B., Auburger G., Bhatnagar S.S., Wastinger B.			
2D	"OPAL1, encoding a dynamn-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28."			
2E	Nat. Genet. 26:211-215(2000).			
2F	-1- FUNCTION: May be involved in mitochondrial biogenesis.			
2G	-1- TISSUE SPECIFICITY: Mitochondrial.			
2H	-1- TISSUE SPECIFICITY: Highly expressed in retina. Also expressed in brain, testis, heart and skeletal muscle.			
2I	-1- DISEASE: Defects in OPAL1 are a cause of optic atrophy type 1 (OPAL1). OPAL1 is a dominantly inherited optic neuropathy occurring in 1 in 50,000 individuals that features progressive loss in visual acuity leading, in many cases, to legal blindness.			
2J	-1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.			
2K	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements).			
2L	OR send an email to license@ebi.ac.uk .			
2M	EMBL, AB011139, BAA25493.1;			
2N	MIM: 605300;			
2O	MIM: 165300;			
2P	InterPro, IPR001401; Dynamn.			
2Q	Protein, P000350; dynamn, 1.			
2R	SMART, PR00195; DYNAMIN.			
2S	SMART, P000033; DINC, 1.			
2T	Motor protein, GTP-binding; Mitochondrion; Transit peptide; Disease mutation; Vison.			
2U	TRANSIT 1			
2V	MITOCHONDRION (POTENTIAL).			
2W	CHAIN 960			
2X	FT_N_BIND 295 302 GTP (POTENTIAL)			
2Y	FT_N_BIND 398 402 GTP (POTENTIAL)			
2Z	FT_N_BIND 467 470 GTP (POTENTIAL).			

Query Match	4.9%	Score 96.5	DB 1	Length 960
Best local similarity	18.2%	Pred. No. 6.4		
Matches 66	Conservative 60	Mismatches 123	Indels 113	Gaps 15
QY	25	NLALKVPELEOCIRELLRLNPNINCICVADLCGASPNITLVVDIYQSIDKYGE	84	R -> Q (IN OPAL). /Frid-Var.011483.
DB	404	NVTYSGMADTKREITSISAKHQPNNILICIDGCVDAKSIYT--DLVSQDPHG--	459	G -> E (IN OPAL). /Frid-Var.011484.
OY	85	EKNLELPTQIIFLN-DLFQNDPNSVFKLDPSTFRKLEKENGKRGISCLISAMPQSYGR	143	MISSING (IN OPAL). /Frid-Var.011485.
DB	460	-----RRTIFVLTKYDLAEKNVAS--PSRIGQI--	203	
OY	144	LPESSMHFLHSCSYVHMLSOVPSGLVIELIGANKSIVSKGCPVQKAYADQFTKD	203	
DB	490	LEPMKALDY-----FAVATKRGNSSEISA--	514	
OY	204	FTTFIRISKELFSGRMILTCICKVDFEDPNPLDLDMAINDLIVEGLLE--EERLDS	261	
DB	515	-----IREEEFFQNSKRLKTSMLKAHQVTRN---LSLAVSDCQWKKVRESVEQOADS	566	
OY	262	FNIPFPTSAEVCVIEEERGSEILYLET--FKAHYDAASIDDDVPAPHSQIKAEVYA	320	
DB	567	FKATFEN-----LETEMKNPNRYRLRELDNRN--ELPEKAKNELID	603	
OY	321	SLIRSEVPEPLASHFGCAIMPDLFHRILKHAAYLH-----MGKCVNINI--ISLAKRP	373	
DB	604	EVIS--LSQVPRHMEBILQOQSLMEVSGVHVIENIYLPAAQWNSGTFTFYTDKLQWT	661	
OY	374	EK 375		
DB	662	DK 663		
RESULT 12				
OPAL_MOUSE				
ID	OPAL_MOUSE	STANDARD	PRT	960 AA.
AC	P58281			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Dynamin-like 120 kDa protein, mitochondrial precursor (large GTP binding protein) (largeG).			
DE	OPAL			
GN	Mus musculus (Mouse).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID:10090			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain.			
RA	Miyashita T., Misaka T., Kubo Y.			
RT	mouse brain.			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
CC	FUNCTION: May be involved in mitochondrial biogenesis.			
CC	1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).			
CC	1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			

or send an email to license@sb-sib.ch.

CC EMBL; AB044138; BAB59000.1;
 CC InterPro; IPR001401; Dynamain.
 CC Pfam; PF00350; dynamin; 1.
 CC PRINTS; PR00195; DYNAMIN.
 CC SMART; SM00053; DYNC1.
 CC Motor protein; gtp-binding; Mitochondrion; Transit peptide.
 CC TRANSIT 1
 CC CHAIN 1
 CC NP_BIND 295 960 DYNAMIN-LIKE 120 KDA PROTEIN.
 CC NP_BIND 302 302 GTP (POTENTIAL).
 CC NP_BIND 398 402 GTP (POTENTIAL).
 CC NP_BIND 467 470 GTP (POTENTIAL).
 CC SEQUENCE 960 AA; 111338 MW; 0F103FB1FD570F49 CRC64;

Query Match 4.9%; Score 96.5; DB 1; Length 960;
 Best Local Similarity 18.2%; Pred. No. 6.4; Mismatches 123; Indels 113; Gaps 15;
 Matches 66; Conservative 60;

CC 25 NLAKVPELEOCIRELLRANLPINCKIRVADLCASGPTLLTVRDIYQIDKVGQE 84;
 CC 404 NTYVSGMAPDTEITFISISKAYMONPNAIILCIODGSVDAERSIYT--DLVSQMDPHG-- 459
 CC 85 EKNELEPTQIQLN--DLFONDENSVEKLLPSEYKLEKENGKIGSCILISAMGSEYGR 143
 CC 460 -----NRITFVLTQYDLAEKVAS-----PSRIQOI-----IEGK 489
 CC 144 LPEPSMHLFHSQSVHMLSGVPSGLVTELGIGANGSIYSSKCGRPYQAKYLQFTKD 203
 CC 490 LFPKALGY-----FAVYTKGNSSESLA----- 514
 CC 204 FTTLRHSELESGRMULTICICKVDEDEPNRDLMLAINDLIVGILE--EKKLDS 261
 CC 515 -----IREYEEFPONSKLTKSMKAKGVTRN-----LSLAVSDCFMKVMESEVDQADS 566
 CC 262 ENIDFETPSAEKVCVEEGSCILLLET--FKAHDAFSTDDIYPRVSEQAKATYA 320
 CC 567 FKATRE-----LETEMKNKPRLEDRN-----ELEKKNKILD 603
 CC 321 SLRSYETPLASHGEAIMPDLFRLAKRAVYLH-----MGKGCYNLI--ISLAKRP 373
 CC 604 EVIS--LSQVTPHWEILIQSLMEVSTHVENIYLPAAQTNGTNGTVDIKLQMT 661
 CC 374 EK 375
 CC 662 DK 663

RESULT 13
 POLG_CXB4E STANDARD; PRT; 2183 AA.
 AC 086887;
 DT 15-JUL-1999 (Rel. 38; Last sequence update)
 DT 15-JUL-1999 (Rel. 38; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DE Genome Polyprotein (Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)).
 OS Cocksackievirus B4 (strain E2).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
 CC NCBI_TaxID=103905;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95205102; Pubmed=7897366;
 RA Kang Y., Chatterjee N.K., Nodwell M.J., Yoon J.W.;
 RT "Complete nucleotide sequence of a strain of coxsackie B4 virus of human origin that induces diabetes in mice and its comparison with nondiabetogenic coxsackie B4 JBV strain."
 RL J. Med. Virol. 44:353-361(1994).

CC -1- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE AT CERTAIN O/G SITES IN THE POLYPEPTIDE. THEY ARE CYSTEINE PROTEASES.
 CC -1- FUNCTION: IT IS THOUGHT THAT THE P3C PROTEIN ATTACHES TO VESICULAR MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- P1A: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/P2A IS CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.
 CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC EMBL; S76772; BAB3885.1; -
 CC HSSP; P21404; 1D4M.
 CC InterPro; IPR004004; Calic1_pol_hel.
 CC InterPro; IPR000199; Cys-protease-3c.
 CC InterPro; IPR003138; Pico_P1A.
 CC InterPro; IPR000081; Pico_P2A.
 CC InterPro; IPR000257; Pico_P2B.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR001205; RNA_helicase.
 CC InterPro; IPR001676; Rny_pol_P3D.
 CC P1am; PF00548; Cys-protease-3c; 1.
 CC P1am; PF02228; Pico_P1A; 1.
 CC P1am; PF00947; Pico_P2A; 1.
 CC P1am; PF01352; Pico_P2B; 1.
 CC P1am; PF00073; Hny; 3.
 CC P1am; PF00680; RNA_deg_RNA_pol; 1.
 CC P1am; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUSNS.
 CC PRODOM; PD001125; Cys-protease-3c; 1.
 CC PRODOM; PD001274; Pico_P2B; 1.
 CC PRODOM; PD001306; Pico_P2A; 1.
 CC Polyprotein; Coat protein; Core protein; Transferase; Myristate;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 CC CHAIN 2
 CC CHAIN 70 330 COAT PROTEIN VP4.
 CC CHAIN 331 568 COAT PROTEIN VP2.
 CC CHAIN 569 846 COAT PROTEIN VP3.
 CC CHAIN 850 999 COAT PROTEIN VP1.
 CC CHAIN 1000 1098 PICORNAIN 2A.
 CC CHAIN 1089 1427 CORE PROTEIN P2B.
 CC CHAIN 1428 1516 CORE PROTEIN P2C.
 CC CHAIN 1517 1538 CORE PROTEIN P3A.
 CC CHAIN 1539 1721 GENOME-LINKED PROTEIN VP6.
 CC CHAIN 1722 2183 PICORNAIN 3C.
 CC LIPID 2
 CC ACT_SITE 1685 1685 MYRISTATE (BY SIMILARITY).
 CC ACT_SITE 1699 1699 PROTEASE (POTENTIAL).
 CC SEQUENCE 2183 AA; 244649 MW; BF44781CF6981D39 CRC64;

Query Match 4.9%; Score 96; DB 1; Length 2183;
 Best Local Similarity 18.3%; Pred. No. 21; Mismatches 154; Indels 108; Gaps 16;
 Matches 73; Conservative 64;

CC 1 MELOEVLMHNEGDSYKNSYNLAKVPELEOCIRELLRANLPINCKIRVADLCG 60
 CC 1816 MRLNAVYGVGEALDITTSAGIPYVLGIK-----RDLSKTKDILKAECKMKY 1869
 CC 61 CASGPTLLTVRDIYQIDKVGQEKNELEPTQIQLNDFONDENSVEKLLPSEYRKL 120
 CC 1870 GLNLP-WTYVKGELRSAEKVAKGRSLTEASSL-----NDSVAMRQTFGNLYKTF 1919

QY 121 EKENGKISGLISAMPSCFYGRFLPESMHFLHSCYSVHMLSGVP-----SGLVIEIG 174
 DB 1920 HANPGLVIGSA-VGCDPDLFWKSIIPMLDGLHRAVDISGASLSPWPAACKLLEKIG 1978
 QY 175 IGANKGSIYSSKGCPRPVQAKYLDQFTKDTFTLRHSHKELFSRGMLLTCLCKVDEDE 234
 DB 1979 -----YSKR-----ETNYIDYLCNSHLY-----RDKHYFVAGKMPSCG----- 2012
 QY 235 PNPDLDMALINDLIVEGL-----EERKLSFNIP----- 265
 DB 2013 -SGTSIFNSMINIITIRLMFKYKGGIDIDOPRMAYGDDVIASYPPLIDASILAAGKG 2071
 QY 266 ---FTPSAEKVICVEEGSC---ELIYLE-TEKHAHYDAFSIDDDVPVSHHQIAEYV 319
 DB 2072 YGLIMTPA-----DKGECNELTNTVTFELKRY---FRADQYPLVHPVMPINDI 2119
 QY 320 ASLIRSVPE-----ILASHGEALIMPD.FHRL 347
 DB 2120 HESITWTDPKNTQYHVSICLLAMHNGEYEEELTPRT 2158

RESULT 14
 POLG_EC30B STANDARD; PRT: 2194 AA.
 ID POLG_EC30B
 AC 09MNT8;
 DT 01-MAR-2002 (Rel. 41, last sequence update)
 DT 01-MAR-2002 (Rel. 41, last annotation update)
 DE Genome polypeptide [contains: Coat protein VP4 (P1a); Coat protein VP2 (P1b); Coat protein VP3 (P1c); Coat protein VP1 (P1d); Picornain 2A (EC 3.4.22.29) (P2a); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPg (P3b); Picornain 3C (EC 3.4.22.28) (P3c); Echovirus 30 (strain Bastian)]
 DE Picornaviruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus
 OC NCBI_taxid=176284;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Zell R.
 RT "Complete genome of echovirus 30 strain Bastian".
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE AT CERTAIN O/G SITES IN THE POLYPEPTIDE. THEY ARE CYSTEINE PROTEASES (BY SIMILARITY).
 CC -1- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis (By similarity).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4 (BY SIMILARITY).
 CC -1- PPM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS. CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC. VP1/P2A IS CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C (BY SIMILARITY).
 CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF162711; AAD45119.1;
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000199; Cys-protease-3c.
 DR InterPro: IPR000138; Pico_P1A.
 DR InterPro: IPR000081; Pico_P2A.
 DR InterPro: IPR002527; Pico_P2B.

DR InterPro: IPR001676; Rhv.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00548; Cys-protease-3c; 1.
 DR Pfam: PF02226; Pico_P1A; 1.
 DR Pfam: PF00947; Pico_P2A; 1.
 DR Pfam: PF01552; Pico_P2B; 1.
 DR Pfam: PF00073; Rhv; 3.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR PRINTS: PR00918; CALICIVIRUSNS.
 DR ProDom: PD001125; Cys-protease-3c; 1.
 DR ProDom: PD001274; Pico_P2B; 1.
 DR ProDom: PD001306; Pico_P2A; 1.
 KW Polypeptide: Coat protein; Core protein; Transferase; Myristate;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 2 69
 FT CHAIN 70 330
 FT CHAIN 331 588
 FT CHAIN 589 860
 FT CHAIN 861 1010
 FT CHAIN 1011 1109
 FT CHAIN 1110 1438
 FT CHAIN 1439 1527
 FT CHAIN 1528 1549
 FT CHAIN 1550 1732
 FT CHAIN 1733 2194
 FT CHAIN 2 2
 FT ACT SITE 1696 1696
 FT ACT SITE 1710 1710
 FT ACT SITE 1710 1710
 SEQUENCE 2194 AA; 245143 MW; F59444CEB091DC9AA CMC64;

Query Match 4.98; Score 96; DB 1; Length 2194;

Best Local Similarity 19.28; Pred. No. 21; Mismatches 156; Indels 90; Gaps 16;

Matches 73; Conservative 61; Mismatches 156; Indels 90; Gaps 16;

QY 1 MELOEVLHNEGEGSDTSTAKNASTNLALAKVPELEOCIRELRLANLPINKICKYADIG 60
 DB 1827 MKLEDAVYGTGELEALDITTSAGYPVALGIRK-----RDLKRTKDLAKLEKMDRY 1880
 QY 61 CASGPTLLTVADIVQSIDKVGQEKNELEPRTIOTFLNDFONDFNSVERKLPLSEYRL 120
 DB 1881 GLNLP-MYVYVDEHRSKAEKAKGKSLIEASL-----NDVAMKOTGNGYKTF 1930
 QY 121 EKENGKISGLISAMPSCFYGRFLPESMHFLHSCYSVHMLSGVP-----SGLVIEIG 174
 DB 1931 HANPGLVIGSA-VGCDPDLFWKSIIPVMDGLHRAVDISGASLSPWPAACKLLEKIG 1989
 QY 175 IGANKGSIYSSKGCPRPVQAKYLDQFTKDTFTLRHSHKELFSRGMLLTCLCKVDEDE 234
 DB 1990 -----YSKR-----ETNYIDYLCNSHLY-----RDKHYFVAGKMPSCG----- 2023
 QY 235 PNPDLDMALINDLIVEGL-----EERKLSFNIPF-----TPSAEYKVICVEEGSC 279
 DB 2024 -SGTSIFNSMINIITIRLMFKYKGGIDIDOPRMAYGDDVIASYPPLIDASILAAGKG 2082
 QY 280 -----EGSC---ELIYLE-TEKHAHYDAFSIDDDVPVSHHQIAEYVSLISYVE 328
 DB 2083 YGLIMTPADKGCPEFNEVMTVTFELKRY---FRADQYPLVHPVMPINDIHSITWD 2139
 QY 329 P-----ILASHGE 337
 DB 2140 PKNTQYHVSICLLAMHNGE 2159

RESULT 15
 POLG_CXA9 STANDARD; PRT: 2201 AA.
 ID POLG_CXA9
 AC P21404;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)

DE Genome polypeptide (contains: Coat protein VP4 (P1A); Coat protein VP2
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
 DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
 OS Coxsackievirus A9 (strain G1993).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 NC NCBL:TaxID=12068;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90111704; PubMed=2558158;
 RA Chang K.H., Aavinen P., Hyypia T., Stanway G.,
 RT The nucleotide sequence of coxsackievirus A9: implications for
 RT receptor binding and enterovirus classification.";
 RL J. Gen. Virol. 70:3269-3280(1989).
 RN [2]
 RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-870.
 RA MEDLINE=20113480; PubMed=10647183;
 RA Hendry E., Hatanaka H., Fry E., Smyth M., Tate J., Stanway G.,
 RA Santti J., Maronen M., Hyypia T., Stuart D.,
 RT "The crystal structure of coxsackievirus A9: new insights into the
 RT uncoating mechanisms of enteroviruses.";
 RL Structure 7:1527-1538(1999).
 CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
 CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VITRO YIELD NATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: D00627; BAA00518.1;
 DR PIR: J00523; GNNVA9.
 DR PDB: 1D4M; 23-DEC-99.
 DR MEROPS: C03.022;
 DR InterPro: IPR000199; Cys-protease-3C.
 DR InterPro: IPR000138; Pico_P1A.
 DR InterPro: IPR000081; Pico_P2A.
 DR InterPro: IPR002527; Pico_P2B.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001676; Rny.
 DR Pfam: PF00548; Cys-protease-3C; 1.
 DR Pfam: PF02226; Pico_P1A; 1.
 DR Pfam: PF00947; Pico_P2A; 1.
 DR Pfam: PF01552; Pico_P2B; 1.
 DR Pfam: PF00073; Rny; 3.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR Pfam: PD001125; Cys-protease-3C; 1.
 DR ProDom: PD001274; Pico_P2B; 1.
 DR ProDom: PD001306; Pico_P2A; 1.
 DR PolyProtein: Coat protein; Core protein; Transferase; 3D-structure;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
 KW CHAIN 2 69 COAT PROTEIN VP4.
 FT CHAIN 330 330 COAT PROTEIN VP2.
 FT CHAIN 331 568 COAT PROTEIN VP3.
 FT CHAIN 569 870 COAT PROTEIN VP1.
 FT CHAIN 871 1017 COAT PROTEIN P2A.
 FT CHAIN 1018 1116 CORE PROTEIN P2B.
 FT CHAIN 1117 1445 CORE PROTEIN P2C.

Query Match	Similarity	Score	DB 1	Length	2201
Best Local	73	Conservative	62	Mismatches	155
Matches	73	Conservative	62	Mismatches	155
1	MELQEVLMNNEGSDTSYAKNSYNLAIAKVPFLBECIRELLRANLPINKICIVADLG 60				
1834	MLEDAVYVTEGLEALDLTTSAGYVVALGIRK-----RDILSKRTROLTLKRCMDKY 1887				
61	CASGPNLTLLVVDIVOSIDKVGGEKNELEPTIOIFLNDLFONDENSVFKLPSFYRKL 120				
1888	GLNLP-MITVYKQDLRSEAKVAKGKSLIEASSL-----NDSVAMRQTFGNLYKTF 1937				
121	EKENGRKIGSCLISAMPSPYGRHPPEESMFLHSCSYVHMLSQVP-----SGLVIELG 174				
1938	HLNPGIVTGA-VQCDPDLFWKRPVLMGLHIAFDYSGYDASISPVWFACLLLEKIG 1996				
175	IGANGSIYSKGCRRPVOKAYIDOPTKOTFTFLRHSKELESGRMLLTCICKYDEDE 234				
1997	-----YSHK-----ETNYIDYLCNSHLX---RDKHYFRGCMPSGC----- 2030				
235	PAPLDLDMALINDIYVGLL-----EEKIKDSFNIPF-----TSAEVCYIYE----- 279				
2031	-SGTIFNSMNNIITFTLMKVKYKIDIDQRMATAGDVIASYPMPIDASLDAKND 2089				
280	-----EEGSC--ETLYLE--TEKAHYDAFSIDDDYVRSHQIKAEYVASLIRSYE 328				
2090	YGLIMTPADKGECEPNEVTWNTVFLKRY---FRADEQYPLVHPVMPKMDIHESLWTKD 2146				
329	P-----ILASHFGE 337				
2147	PNTQDHYRSCLLAWHNGE 2166				

Search completed: July 27, 2002, 05:47:03
 Job time: 414 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2002, 05:34:04 : Search time 62.29 Seconds

(without alignments)
1049.801 Million cell updates/sec

Title: US-09-971-020-1

Sequence: 1 MELOEVLHNEEGEDTSYAK.....KGCYNIIISLAKPKSDV 378

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:19:1
2: SP_BACTERIA:1
3: SP_FUNGI:1
4: SP_HUMAN:1
5: SP_INVERTEBRATE:1
6: SP_MAMMAL:1
7: SP_MHC:1
8: SP_ORGANELLE:1
9: SP_PHAGE:1
10: SP_PLANT:1
11: SP_PROTOZOA:1
12: SP_VIRUS:1
13: SP_VERTEBRATE:1
14: SP_UNCLASSIFIED:1
15: SP_XVIRUS:1
16: SP_BACTERIAP:1
17: SP_ARCHAEP:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1972	100.0	378	10	Q9AVJ9 coffee arab
2	1632	82.8	372	10	Q9AVK0 coffee arab
3	1593.5	80.8	385	10	Q9AVK1 coffee arab
4	1569.5	79.6	385	10	Q9AVL9 coffee arab
5	723.5	36.7	366	10	Q9AVR1 stephanotis
6	689	34.9	357	10	Q9AVR1 atropa bell
7	662.5	33.6	352	10	Q9SBR6 brassica ra
8	649.5	33.0	359	10	Q9SPV4 clarkia bre
9	649.5	32.9	364	10	Q9FYV9 antirrhinum
10	639	32.4	389	10	Q9AR07 arabidopsis
11	635.5	32.2	388	10	Q9LRU5 arabidopsis
12	633	32.1	384	10	Q9FRT6 arabidopsis
13	629	31.8	383	10	Q9ISZ0 arabidopsis
14	628	31.8	389	10	Q9FZM8 camelina sat
15	614.5	29.7	371	10	Q23234 arabidopsis
16	586	29.7	327	10	Q9CAX3 arabidopsis

17	577.5	29.3	359	10	Q92PR3 arabidopsis
18	526.5	26.7	378	10	Q943A3 oryza sativ
19	523	26.5	323	10	Q9FMR8 arabidopsis
20	459	23.3	380	10	Q9FWJ1 oryza sativ
21	451	22.9	351	10	Q93XW1 brassica ca
22	445.5	22.6	352	10	Q9KPD0 arabidopsis
23	439	22.3	351	10	Q9K157 arabidopsis
24	432	21.9	386	10	Q9FLW8 arabidopsis
25	425	21.6	359	10	Q9LS10 arabidopsis
26	423.5	21.5	353	10	Q9C9W8 arabidopsis
27	416.5	21.1	374	10	Q9FKC8 arabidopsis
28	413.5	21.0	352	10	Q9C9M4 arabidopsis
29	413.5	20.9	619	10	Q9S592 arabidopsis
30	404.5	20.5	379	10	Q9FYC3 arabidopsis
31	401	20.3	348	10	Q9FYC4 arabidopsis
32	400	20.3	348	10	Q9FYC6 arabidopsis
33	396	20.1	348	10	Q949P9 arabidopsis
34	392.5	19.9	353	10	Q9M4F2 brassica na
35	392	19.9	353	10	Q9C9M3 arabidopsis
36	387.5	19.7	353	10	Q93YL7 brassica na
37	387	19.6	353	10	Q9C9M2 arabidopsis
38	385	19.5	361	10	Q9FKR0 arabidopsis
39	373.5	18.9	361	10	Q9FMA2 arabidopsis
40	360	18.3	318	10	Q9FTJ3 arabidopsis
41	198.5	10.1	210	10	Q941F2 arabidopsis
42	127.5	6.5	534	16	Q98QW7 mycoplasma
43	107.5	6.0	438	3	Q05919 saccharomyc
44	107.5	5.5	999	17	Q9Y914 aecorytum p
45	106.5	5.4	585	4	Q965X5 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	378 AA.
ID	Q9AVJ9			
AC	Q9AVJ9			
DT	01-JUN-2001 (TREMUR17, 17, Created)			
DT	01-JUN-2001 (TREMUR17, 17, Last sequence update)			
DT	01-DEC-2001 (TREMUR19, 19, Last annotation update)			
DE	7-METHYLANTHINE N-METHYLTRANSFERASE.			
GN	CAMANT.			
OS	Coffea arabica (Coffee).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridiales; euasterids I; Gentianales; Rubiaceae; Ixoriaceae; Coffeae;			
OC	Coffea.			
OX	NCBI_TaxID-13443;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. CATURRA;			
RX	MEDLINE-2169383; PubMed-11108716;			
RA	Ogawa M, Heral Y, Koizumi N, Kusano T, Sano H;			
RT	7-METHYLANTHINE METHYLTRANSFERASE OF COFFEE PLANTS. GENE ISOLATION			
RT	AND ENZYMATIC PROPERTIES.			
RL	J. Biol. Chem. 276:8213-8218(2001).			
DR	EMBL: AB046794; BAB9226.1;			
KW	Transferase; Methyltransferase.			
SQ	SEQUENCE 378 AA; 42747 MW; F36184F3D5243809 CRC64;			

Query Match 100.0%; Score 1972; DB 10; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.1e-161;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	MELOEVLHNEEGEDTSYAKNAYSLAKVPPLEOCIRELRLANPININKIKYADIG 60
DB	1	MELOEVLHNEEGEDTSYAKNAYSLAKVPPLEOCIRELRLANPININKIKYADIG 60
OY	61	CAGPRTLTLYDIOGSDKVGGEKRELEPTIOFLINDLFQNDPNSVKLLPSYRKL 120
DB	61	CAGPRTLTLYDIOGSDKVGGEKRELEPTIOFLINDLFQNDPNSVKLLPSYRKL 120

QY 121 EKENGRKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
DB 121 EKENGRKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
QY 181 SIYSSKGRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
DB 181 SIYSSKGRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
QY 241 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 300
DB 241 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 300
QY 301 SIDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 360
DB 301 SIDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 360
QY 361 CYNMLIISLAKKPEKSDV 378
DB 361 CYNMLIISLAKKPEKSDV 378

RESULT 2

QYAVKO PRELIMINARY; PRT; 372 AA.
AC QYAVKO;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE THEOBROMINE SYNTHASE.
GN CAMT13.
OS Coffea arabica (Coffee).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Rubiaceae; Ixoroideae; Coffeae;
OC Coffea.
OX NCBI_TaxID=13443;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CY. CATURRA;
RX MEDLINE=21269383; PubMed=11108716;
RA Ogawa M., Herai Y., Koizumi N., Kusano T., Sano H.;
RT *7-Methylxanthine Methyltransferase of Coffee Plants. GENE ISOLATION
AND ENZYMAIC PROPERTIES.
RL J. Biol. Chem. 276:8213-8218(2001).
DR EMBL; AB048793; BAB39215.1;
SQ SEQUENCE 372 AA; 41842 MW; 1C9D2D53263828D2 CRC64;

Query Match 82.8%; Score 1632; DB 10; Length 372;
Best Local Similarity 84.7%; Pred. No. 3.5e-132;
Matches 321; Conservative 16; Mismatches 34; Indels 8; Gaps 2;

QY 1 MELOEVYHMHNEGEGDTSYAKNASYN-LALAKVRFLEOCIRELRLANLPINIKCIVADL 59
DB 1 MELOEVYHMHNEGEGDTSYAKNASYN-LALAKVRFLEOCIRELRLANLPINIKCIVADL 59
QY 60 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 119
DB 60 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 119
QY 61 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 120
DB 61 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 120
QY 120 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 179
DB 120 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 179
QY 121 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
DB 121 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
QY 180 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 239
DB 180 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 239
QY 181 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
DB 181 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
QY 240 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 299
DB 240 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 299
QY 361 CYNMLIISLAKKPEKSDV 378
DB 361 CYNMLIISLAKKPEKSDV 378

QY 300 FSIDDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 359
DB 300 FSIDDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 359
QY 360 CYNMLIISLAKKPEKSDV 378
DB 360 CYNMLIISLAKKPEKSDV 378

RESULT 3

QYAVKO PRELIMINARY; PRT; 385 AA.
AC QYAVKO;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE THEOBROMINE SYNTHASE.
GN CAMT13.
OS Coffea arabica (Coffee).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Rubiaceae; Ixoroideae; Coffeae;
OC Coffea.
OX NCBI_TaxID=13443;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CY. CATURRA;
RX MEDLINE=21269383; PubMed=11108716;
RA Ogawa M., Herai Y., Koizumi N., Kusano T., Sano H.;
RT *7-Methylxanthine Methyltransferase of Coffee Plants. GENE ISOLATION
AND ENZYMAIC PROPERTIES.
RL J. Biol. Chem. 276:8213-8218(2001).
DR EMBL; AB048793; BAB39215.1;
SQ SEQUENCE 385 AA; 43270 MW; 92103A20A001FBOE CRC64;

Query Match 80.8%; Score 1593.5; DB 10; Length 385;
Best Local Similarity 81.3%; Pred. No. 7.6e-129;
Matches 313; Conservative 25; Mismatches 40; Indels 7; Gaps 2;

QY 1 MELOEVYHMHNEGEGDTSYAKNASYN-LALAKVRFLEOCIRELRLANLPINIKCIVADL 59
DB 1 MELOEVYHMHNEGEGDTSYAKNASYN-LALAKVRFLEOCIRELRLANLPINIKCIVADL 59
QY 60 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 119
DB 60 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 119
QY 61 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 120
DB 61 GCASGPTLLTVADIVOSIDKVGGEKNELEPPTIOIFLNDLFONDENSVFKLLPSFYRK 120
QY 120 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 179
DB 120 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 179
QY 121 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
DB 121 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
QY 180 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 239
DB 180 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 239
QY 181 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
DB 181 GSISYSGCRPPVOKAYLDQFTKDFTEFLRIHSEKELFSRGRLMLTICICVDEFPDNPDL 240
QY 240 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 299
DB 240 LDMALNDLIYEGLEBEERKIDSFNIPFTPSAEVKCIVEEGSCILYLETFFKAHYDAF 299
QY 300 FSIDDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 353
DB 300 FSIDDDYVPSHEDQIRAEVYASLIRSYVEPIILASHGEALIMPOLFHRLAKHAKVILHNGK 353
QY 360 CYNMLIISLAKKPEKSDV 378
DB 360 CYNMLIISLAKKPEKSDV 378
QY 361 CYNMLIISLAKKPEKSDV 378
DB 361 CYNMLIISLAKKPEKSDV 378

RESULT 4
QYAVKO PRELIMINARY; PRT; 385 AA.
AC QYAVKO;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE THEOBROMINE SYNTHASE.
GN CAMT13.
OS Coffea arabica (Coffee).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Rubiaceae; Ixoroideae; Coffeae;
OC Coffea.
OX NCBI_TaxID=13443;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CY. CATURRA;
RX MEDLINE=21269383; PubMed=11108716;
RA Ogawa M., Herai Y., Koizumi N., Kusano T., Sano H.;
RT *7-Methylxanthine Methyltransferase of Coffee Plants. GENE ISOLATION
AND ENZYMAIC PROPERTIES.
RL J. Biol. Chem. 276:8213-8218(2001).
DR EMBL; AB048793; BAB39215.1;
SQ SEQUENCE 385 AA; 43270 MW; 92103A20A001FBOE CRC64;

DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE CAFFEINE SYNTHASE.
 GN CAMELLIA.
 OS Coffea arabica (Coffee).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Rubiaceae; Ixoroidae; Coffeae;
 OC Coffea.
 NCBI_TaxID=13443;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CALLUS; PubMed=11108716;
 RA Ogawa M., Heral Y., Koizumi N., Kusano T., Sano H.;
 RT "7-Methylxanthine Methyltransferase of Coffee Plants. GENE ISOLATION
 RT AND ENZYMATIC PROPERTIES."
 RL J. Biol. Chem. 276:8213-8218(2001).
 DR EMBL; AB039725; BAB39213.1;
 SO SEQUENCE 385 AA; 43242 MW; 003A5837F0AE300 CRC64;

Query Match 79.6%; Score 1569.5; DB 10; Length 385;
 Best Local Similarity 80.3%; Pred. No. 8; Be-127;
 Matches 309; Conservative 26; Mismatches 43; Indels 7; Gaps 2;

OY 1 MELOEVLHNNGBEDTSTAKNASTN-LALAKYKPLEDCIRELLRANLPINCKIKYADL 59
 DB 1 MELOEVLHNNGBEDTSTAKNASTN-LALAKYKPLEDCIRELLRANLPINCKIKYADL 60
 OY 60 GCASGPNLTLVNDIVOSIDKVGQEKNELEPTQIFLNDLFONDNSVFKLLPSFYRK 119
 DB 61 GCASGPNLTLVNDIVOSIDKVGQEKNELEPTQIFLNDLFONDNSVFKLLPSFYRK 120
 OY 120 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 179
 DB 121 LEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELGIGANK 180
 OY 180 GSIVSSKCRPPVOKAYLDOFTKDTFTLRHSEKLEFRGMILTCCKYDEDEPNPLD 239
 DB 181 GSIVSSKCRPPVOKAYLDOFTKDTFTLRHSEKLEFRGMILTCCKYDEDEPNPLD 240
 OY 240 LLDMAINDLIVGELLEBEKIDSFNIFPTSPAEKCIYEESGCELTILETFKAHYDAA 299
 DB 241 LLDMAINDLIVGELLEBEKIDSFNIFPTSPAEKCIYEESGCELTILETFKAHYDAA 300
 OY 300 FSIDDDYVRSH-----EQIKAEVYASLIRSVYEPILASHFGAIMPDLFRLAKNAK 353
 DB 301 FSIDDDYVRSH-----EQIKAEVYASLIRSVYEPILASHFGAIMPDLFRLAKNAK 360
 OY 354 VLMGKGCYNNLIISLAKKPEKSDV 378
 DB 361 VIRLKGYYNLIISLAKKPEKSDI 385

RESULT 5
 O9AVR1 PRELIMINARY; PRT; 366 AA.
 AC O9AVR1.
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE S-ADENOSYL-L-METHIONINE: SALICYLIC ACID CARBOXYL
 DE METHYLTRANSFERASE.
 GN SAMT.
 OS Stephanotis floribunda.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Apocynaceae; Asclepiadoideae;
 OC Ceropapeeae; Stephanotis.
 NCBI_TaxID=85838;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Pott M.B., Pichersky E., Piechulla B.;
 RT "Circadian oscillation of methyl salicylate emission, samt enzyme
 RT activity, and samt mRNA in flowers of Stephanotis floribunda."
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ308570; CAC33768.1;
 SO SEQUENCE 366 AA; 41316 MW; 8D5676AB153F88BD CRC64;

Query Match 36.7%; Score 723.5; DB 10; Length 366;
 Best Local Similarity 42.7%; Pred. No. 4; 7e-54;
 Matches 164; Conservative 64; Mismatches 125; Indels 31; Gaps 10;

OY 1 MELOEVLHNNGBEDTSTAKNASTN-LALAKYKPLEDCIRELLRANLPINCKIKY 56
 DB 1 MELOEVLHNNGBEDTSTAKNASTN-LALAKYKPLEDCIRELLRANLPINCKIKY 54
 OY 57 ADLGCASGPNLTLVNDIVOSIDKVGQEKNELEPTQIFLNDLFONDNSVFKLLP 114
 DB 58 ADLGCASGPNLTLVNDIVOSIDKVGQEKNELEPTQIFLNDLFONDNSVFKLLP 110
 OY 115 SFYRKLEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELG 174
 DB 111 SFYRKLEKENGRIKISGLISAMPSEFYGRLEPESMHFLHSCSYVHMLSOVPSGLVTELG 165
 OY 175 IGANKSIYSSKCRPPVOKAYLDOFTKDTFTLRHSEKLEFRGMILTCCKYDEDEPNPLD 231
 DB 166 -EVNKGNIITLSTPSVIRAYLQFORDTTLQCRABEILPGVAVLTMGKKGDS 224
 OY 232 -FDEPNLDDMAINDLIVGELLEBEKIDSFNIFPTSPAEKCIYEESGCELTILE 290
 DB 225 GRESGYALELTLRALNELVSGLEEQDLCFVNPQYTPSPAEVYFEEESGSTRLE 284
 OY 291 TPKAHYDAFSDIDDYVRSHDOIK--AEVYASLIRSVYEPILASHFGAIMPDLFRLA 348
 DB 285 ATTHTHTA--VDHGHVGHGHHAFKDCGYSLSNCRVAVPELVHFGAALMDEVFHYR 341
 OY 349 KHAAYVLMGKGCYNNLIISLAKK 372
 DB 342 EILNCTMEKIEFIVNVSMKR 365

RESULT 6
 O9AVG9 PRELIMINARY; PRT; 357 AA.
 AC O9AVG9.
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE S-ADENOSYL-L-METHIONINE: SALICYLIC ACID CARBOXYL
 DE METHYLTRANSFERASE.
 GN ABSAMT1.
 OS Atropa belladonna (Belladonna).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.
 NCBI_TaxID=33113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M8; TISSUE=ROOT;
 RA Fukami H., Asakura T., Hirano H., Abe K., Shimomura K., Yamakawa T.;
 RT "Cloning and expression of salicylic acid inducible and active S-
 RT adenosyl-L-methionine:salicylic acid carboxyl methyltransferase in
 RT transformed root culture of Atropa belladonna."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB049752; BAB39396.1;
 SO SEQUENCE 357 AA; 40678 MW; FA6A1738FA967EF2 CRC64;

Query Match 34.9%; Score 689; DB 10; Length 357;
 Best Local Similarity 42.0%; Pred. No. 4; 2e-51;
 Matches 161; Conservative 68; Mismatches 116; Indels 38; Gaps 10;

```

OY 1 MELOEVLHMHNEGEGTSTAKNASTINLAKV---KPELEOCIRELLRANLPINCKIRVADL 56
DB 1 MKVYEVLMHNGNGNGISTANN---SLVORVILMTKPTTEQALIDLVCSFPF---ETLCI 54
OY 57 ADGCGAGNTLLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 116
DB 55 ADGCGAGNTLLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 112
OY 117 YKLEKENGKRTIGSCLISAMPSPFYGRLEPESMHFLHSCYVHMLSOVPSGLVIELGIG 176
DB 113 QODLRQJGEEGPEFSPGSPFYGRLEPESMHFLHSCYVHMLSOVPSGLVIELGIG 166
OY 177 ANKGSITSSKCCRPVOKAYLDQFTDFTFLRIHSEKLEFSGRMILITCICKVDEDEPN 236
DB 167 KKNENITASTSPSPVAKAYKQEKDFNFLEKSELMKGKMYLTFELGR---ESEDPS 224
OY 237 P-----LDLDAINDLIVEGLLEEKLDSPNIPFTPSAEVKCIVEEGSCILYLE 290
DB 225 SECCITWELLSMALNELVLEGLIEEEKVDSPNIOYTPSPPEEKYIVREGSFTINRLE 284
OY 291 TKRAYDAAFSIDDDYVNSHQIKAEY-VASLIRSYEPLASHGEAIMPDLEFRLAK 349
DB 285 ATRVHMV-----SNEGINGVAVAKOMAAVEPLVSOFPDOKLMLVFOKYE 333
OY 350 HAAKVLHMGKGCYNLITSLAK 372
DB 334 IISDISEKTEFETINVIVLSLTK 356

RESULT 7
OY 09SRK6 PRELIMINARY; PRT; 392 AA.
AC 09SRK6:
DT 01-MAY-2000 (TREMBLER, 13, Created)
DT 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE FLOAL NECTARY-SPECIFIC PROTEIN.
GN NTR1.
OS Brassica rapa subsp. pekinensis (Chinese cabbage) (Celeriac cabbage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eucosids II; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX NCBI_TaxID=51351;
RN [1].
RP MEDLINE-9306827; PubMed-10809010;
RA Song J.T., Seo H.S., Song S.I., Lee J.S., Choi Y.D.;
RT "NMR encodes a floral nectary-specific gene in Brassica campestris L.
RT plant Mol. Biol. 42:647-655(2000).
DR EMBL: AF179222; AAF22289.1;
SQ SEQUENCE 392 AA; 43815 MW; 25B78530E93B5757 CRC64;

Query Match 33.6%; Score 662.5; DB 10; Length 392;
Best Local Similarity 37.6%; Pred. No. 9.1e-49;
Matches 151; Conservative 77; Mismatches 131; Indels 43; Gaps 11;

OY 1 MELOEVLHMHNEGEGTSTAKNASTINLAKV---KPELEOCIRELLRANLPINCKIRVADL 59
DB 1 MEVRIHMHKNGGEGTSTAKNASTINLAKV---KPELEOCIRELLRANLPINCKIRVADL 58
OY 60 GCASGPNLTLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 117
DB 59 GCASGPNLTLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 114
OY 118 RKLKLEK-----ENKRTIGSCLISAMPSPFYGRLEPESMHFLHSCYVHMLSOVPSGLVIELGIG 166
DB 115 DRKAKDNVYESLGEHSGG-GPCFSAVAGSYGRLFRRLRSLHVSLSLHMLSOVPC 173
OY 167 -----SGVIELGIGANKGSITSSKCCRPVOKAYLDQFTDFTFLRIHSEKLEFSGRMILITCICKVDEDEPN 220
DB 174 GEVNMKDDGVITADLD-NKGNITLTKSPSAKAKYALQFOTDPSVPLRSLRSEELVPGGR 232

```

```

OY 221 MLTNCICKVDEDEPNPL-----LDLDAINDLIVEGLLEEKLDSPNIPFTPSAE 272
DB 223 MVLSFLGR-----SSDPPTTESCYOMELQALSLAKGILIEENIDAFNAPYLAASPE 288
OY 273 EVKCIYEEGSCGELILEFRAHDAAFSIDDDYV---RSHEQIKAEYVASLIRSYEP 329
DB 289 ELKMAIEKESFSIDRLSEISYVDWEGSISDDSDIVYKRPALASGRRAKTRAVAYER 348
OY 330 ILASHGEAIMPDLEFRLAKAAVLMKGCYNLITSLAK 371
DB 349 MLERTFOQVMDLEFERYAKLVGEIVYSSPRITIVYSLR 390

RESULT 8
OY 09SPV4 PRELIMINARY; PRT; 359 AA.
AC 09SPV4:
DT 01-MAY-2000 (TREMBLER, 13, Created)
DT 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE S-ADENOSYL-L-METHIONINE: SALICYLIC ACID CARBOXYL
DE METHYLTRANSFERASE.
GN SAMT.
OS Clarkia breweri.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=36903;
RN [1].
RP SEQUENCE FROM N.A.
RA MEDLINE-9306827; PubMed-10375393;
RA Ross J.R., Nam K.H., D'Auria J.C.
RT "S-Adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, an
RT enzyme involved in floral scent production and plant defense,
RT Arch. Biochem. Biophys. 367:9-16(1999).
DR EMBL: AF13053; AAF00108.1;
SQ SEQUENCE 359 AA; 40289 MW; B0FE3E41MBBD851 CRC64;

Query Match 33.0%; Score 651.5; DB 10; Length 359;
Best Local Similarity 37.9%; Pred. No. 7.1e-48;
Matches 148; Conservative 75; Mismatches 119; Indels 49; Gaps 11;

OY 1 MELOEVLHMHNEGEGTSTAKNASTINLAKV---KPELEOCIRELLRANLPINCKIRVADL 59
DB 1 MDVROVLHMKGAGENSYANSPIOQVISTIKPTTEAITALYSGD--TVTTRIALDL 58
OY 60 GCASGPNLTLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 115
DB 59 GCASGPNLTLVRIYDIOVSDKVGQEKNELEPTIOIFLNDLFONDENVFKLLPSEF 112
OY 116 FYRKLKENGKRTIGSCLISAMPSPFYGRLEPESMHFLHSCYVHMLSOVPSGLVIELGIG 175
DB 113 -----TENDVD---GVCINGVPOSGFYGRLEPESMHFLHSCYVHMLSOVPSGLVIELGIG 159
OY 176 GANKSITSSKCCRPVOKAYLDQFTDFTFLRIHSEKLEFSGRMILITCICKVDEDEPN 235
DB 160 ESNKGNITMANTPOGSLANAYKQFQEDHALFLCRQDQVYVPGRAVLITILGRSE-DRA 218
OY 236 NP-----LDLDAINDLIVEGLLEEKLDSPNIPFTPSAEVKCIVEEGSCILYLE 290
DB 219 STEOCLIMOLLAMALNOMVSEGLIEEKMDKFNIPOTTPTEVAMEILKESGFLIDHIE 278
OY 291 TKRAYH-----DAFSIDDDYVNSHQIKAEYVASLIRSYEPLASHGEAIMPDLEFRLAK 343
DB 279 ASEIYSSCTKDDGGGSVEE-----GYNVARKRAVAEPLLDHFGAILIEDV 328
OY 344 FHLRLAKAAVLMKGCYNLITSLAK 374
DB 329 FHRVYKILITIERMSKEKTEKFINIVYVLSLRKSD 359

```

```

RESULT 9
ID 09FY29 PRELIMINARY: PRT: 364 AA.
AC 09FY29:
DT 01-MAR-2001 (TREMblrel, 16, Created)
DT 01-MAR-2001 (TREMblrel, 16, Last sequence update)
DE 01-JUN-2001 (TREMblrel, 17, Last annotation update)
DE SAM-BENZOIC ACID CARBOXYL METHYLTRANSFERASE.
GN BMT.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20312758; PubMed=10852939;
RA Dudareva N., Murfitt L.M., Mann C.J., Gorenstein N., Kolosova N.,
RA Kish C.M., Bonham C., Wood K.;
RT "Developmental Regulation of Methyl Benzoate Biosynthesis and Emission
RT in Snapdragon Flowers."
RT Plant Cell 12:949-961(2000).
RT EMBL: AF198492; AAF98284.1;
DR InterPro: IPR002106; AA.TRNA.LIGASE.11;
DR PROSITE: PS00339; AA.TRNA.LIGASE.11.2; UNKNOWN.1.
KM Transferase: Methyltransferase.
SQ SEQUENCE 364 AA; 41010 MW; C10C8E864A581419 CRC64;

Query Match 32.4%; Score 649.5; DB 10; Length 364;
Best Local Similarity 37.5%; Pred. No. 1.1e-47;
Matches 143; Conservative 85; Mismatches 122; Indels 31; Gaps 10;

QY 3 LQEVLMHN-EGEDPTSYAKNAS-YNLALAKV-PLEECIRELLRANLPINIKCIKIVADLG 60
DB 4 MKKLCLMNGIDGETSYANNSSGLOKYMMSKSLHVDLTKIDIGDHY-GFPCCFMDMG 62
QY 61 CASGPNLTIVROIVOSIDKVGQGEKNLEPRPTIOIFLNDLFONDFNSVFKLIPFVYKL 120
DB 63 CSQGNALVAVSGIYMTIEDL-YTERKININELPEFEVELNDLPNDENNFKL-----L 114
QY 121 EKENRRTGSGILSAMGSGYGRLPPEESMHPLSCSVHMLGOVSGVIELGIGANKG 180
DB 115 SHEN-----GNCFYVGLPGSFYGRLLPKKSLHFAVSSYIMLSQVEGLE-----DNKRQ 165
QY 181 SYSSKGRPPVOKAYLDOTKDTFTFLRHSKELFSGRMULTICICKYDEDERNPID- 239
DB 166 NYKMTESPPEVTKAIKAKOTERDPSTFLKREELVPEGRVLEFNGR--SVEDPSKOD 223
QY 240 -----LDMALINDIVLVEGLEEKLDSEFNIPPTPSAEVYKCVIEEGSCETILYETFA 294
DB 224 IAFITLAKTLVMAVAGVKMDLYSFNIPYSPCTREVEAAILSEGSFTLDRELEFRV 283
QY 295 HYDAFSDIDD--YVRSHEQIKAEYVASLRSYEPITLASHGGAIMPDLFHLAKHA 351
DB 284 CWDADYDDDDDDDDPSIFGKQSGKGFVADCAVRAITEPMILASHGSGTIDMLFFGYAKKI 343
QY 352 AKVLMGKGCYNNLIISLAK 372
DB 344 VEHLSVENSYSFISVLSLR 364

RESULT 10
ID 09AR07 PRELIMINARY: PRT: 389 AA.
AC 09AR07:
DT 01-JUN-2001 (TREMblrel, 17, Created)
DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel, 17, Last annotation update)
DE S-ADENOSYL-L-METHIONINE:JASMONIC ACID CARBOXYL METHYLTRANSFERASE.
GN JMT.
OS Arabidopsis thaliana (Mouse-ear cress).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11287657;
RA Seo H.S., Song J.T., Cheong J.J., Lee Y.H., Lee Y.W., Hwang I.,
RA Lee J.S., Choi Y.D.;
RT "Jasmonic acid carboxyl methyltransferase: A key enzyme for jasmonate-
RT regulated plant responses."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4788-4793(2001).
RL EMBL: AY008435; AAC23344.1;
DR EMBL: AY008434; AAC23343.1;
SQ SEQUENCE 389 AA; 43372 MW; 75A819E77662A18E CRC64;

Query Match 32.4%; Score 639; DB 10; Length 389;
Best Local Similarity 37.8%; Pred. No. 9.4e-47;
Matches 151; Conservative 79; Mismatches 129; Indels 40; Gaps 12;

QY 1 MELQEVLMHNEGEDPTSYAKNASYNLALAKV-KPLEECIRELLRANLPINIKCIKIVADL 59
DB 1 MEVRYRLMKNNGNGETSYAKNSTAQSNIISLGRVMDALTKLMNS--SEISSIGIDL 58
QY 60 GCSGPNLTIVROIVOSIDKVGQGEKNLEPRPTIOIFLNDLFONDFNSVFKLIPFVYKL 117
DB 59 GCSGPNLSLISNIVDTINLCP-----DLDRPELVSLNDLPSPDFYICASLPFVY 114
QY 118 RLKLEKNGRKIG-----SCLISAMGSGYGRLPPEESMHPLSCSVHMLGOV-- 166
DB 115 --DRYNNKRELGFGRGGECSVSAVDSFGRLPRSLHVAHSSSLHMLSVPCR 171
QY 167 -----SGVIELGIGANKSGSYSSKGRPPVOKAYLDOTKDTFTFLRHSKELFSGR 220
DB 172 EAEREDRTTADL--ENMGRIYIKTSRPSAKKAYALOTDVLVLRSSSELVPGGR 228
QY 221 MLTCTICK--VDEDEPN--PLDLMALINDIVLVEGLEEKLDSEFNIPPTPSAEVYK 276
DB 229 MYSEFGRSRLDPTTEESCYOMELLAQALMSNAKEGIIIEEKIDAFNAPYAAASSELKM 288
QY 277 IVEEGSCETILYETFAHYDAFSDIDDV--YVRSHEQIKAEYVASLRSYEPITL 332
DB 289 VIKREGSFSDIRLEISPIDWEGSISESYDLAIRKPEALASGRVSNIRAVVEPML 348
QY 333 SHGGAIMPDLFHLAKHAAYLMGKGCYNNLIISLAK 371
DB 349 PTGENVMDLPERIAKIVGEYFVSSPRATVILSVR 387

RESULT 11
ID 09LR15 PRELIMINARY: PRT: 368 AA.
AC 09LR15:
DT 01-OCT-2000 (TREMblrel, 15, Created)
DT 01-OCT-2000 (TREMblrel, 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel, 19, Last annotation update)
DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE-LIKE
DE PROTEIN. FLORAL NECTARY-SPECIFIC PROTEIN-LIKE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX SPRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asanizu E., Tabata S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;

```


QY 127 KISSCLISAMSGSFYGRPFESNMFLHSCYSVHWLSQVPSGLVIELGIGANKSIYSSK 186
 DB 113 VKKCKISVPSGFSKSLPESKSLHFWSSLCIMWLSKYPDL-----EDNKKYLNLS 166
 QY 187 GCBPPOKAYLDOFTKDTTFLRIHSEKELFSRGMMLTICICKVDEFPDPLDLM--- 243
 DB 167 PCBPYKSKYLQFKNDSFLRLRADEYVPGKRMALTFVGR-----KSLDPLSKDCP 219
 QY 244 -----AINDIYEGLEEEKLDSFNIFPFPSSAEYKCIIVEEGSCETILYETPKAH 295
 DB 220 QNMSISDLDLVSEGIYKSDVDNSNLFNIPDSEVREYIESEGSFKISNFTI--- 276
 QY 296 YDAFSEI-----DDDYVRSHEQIKAEYV--ASLRSVPEPLIASHFGGAIIMPDL 343
 DB 277 FGLLFYKTKRGREYKDDDD---NLDOSCFEYIKRKASLINSITEPMLGAFGDALMDRL 333
 QY 344 FHLAKHAK 353
 DB 334 FERYTHLAE 343

RESULT 14
 ID 09FZNB PRELIMINARY; PRT: 369 AA.
 AC 09FZNB;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CAFFEINE SYNTHASE.
 GN TCS1.
 OS Camellia sinensis (Tea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Ericales; Theaceae; Camellia.
 OX NCBI_TaxID=4442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20437355; PubMed=10984041;
 RA Kato M., Mizuno K., Crozier A., Fujimura T., Ashihara H.;
 RL "Caffeine synthase gene from tea leaves".
 DR EMBL; AB031280; BAB12278.1;
 SO SEQUENCE 369 AA; 41272 MW; E6D262087E475576 CRC64;

Query Match 31.8%; Score 628; DB 10; Length 369;
 Best Local Similarity 36.3%; Pred. No. 7,8e-46;
 Matches 143; Conservative 80; Mismatches 113; Indels 56; Gaps 11;
 QY 2 ELQEVLMHNEGEGSTYAKNASTNLAKV KPLFEOCIRELLRANLPINCKIRVADLG 60
 DB 8 KYNEVLFMRKGESESTYAKNSFTQOVASMAOPALENAVETLESRDH--LQALNADLG 65
 QY 61 CASGPNLTLLTVRDIVOSIDKVGQEEKNELERT--IQIFLNDLFQDNFNSVFLLPSFYRL 118
 DB 66 CAGAGNHTA-----VISTIKRMRKRCRELNCOTLELOYLLDNGDNFNLFGLSS--- 118
 QY 119 KLEKENGKTKS-----CLISAMPSFGRLFPESNMFLHSCYSVHWLSQVPSGLVIE 172
 DB 119 -----EVINKKEEPPCYVMGVPSEFGRLEFPNSLHSLVSHVWLSQVPSGLVIE 171
 QY 173 LGTIGANKSIYSSGCRPPVOKAYLDOFTKDTTFLRIHSEKELFSRGMMLTICICKVDEF 232
 DB 172 EGIANKRKIIISTSPYVKNLSQFHEDTFLMARSQEVVPCNWL--TLNGRC 229
 QY 233 DEPNL-----DLIDMAINDLIVEGLEEEKLDSFNIPFPSPAEVAVCIYEEBSCETI 286
 DB 230 SPSDMQSCFWELLMALMAYLSQGLIDEDKLDFTNIPSPFASLEVKDIYERDGSFTI 289
 QY 287 LYLETFKHYAASIDDOVPVRSHEQIKAEYVASLRSVPEPLIASHFGGAIIMPDLFHR 346
 DB 290 DHE-----GPDLDVSVMQENDKRWAGEKTYVNAFTFELISNQSPFELMDK 341

QY 347 LAKHAAVLMHGKCYNNLIIS--LAKKPEKSDV 378
 DB 342 -----FTHIYVSLDEAKILPKTTSI 360
 RESULT 15
 ID 023234 PRELIMINARY; PRT: 371 AA.
 AC 023234;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYDROTHERMAL 42.0 KDA PROTEIN.
 GN C7A10.890 OR AT4G36470.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Terry N., Vos P., Heljnen L., Mewes H.W., Schueller C.,
 RA Chaitwalz N.,
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; Z99708; CAB16845.1;
 DR EMBL; AL161589; CAB80313.1;
 KW Hypothetical protein.
 SO SEQUENCE 371 AA; 41984 MW; F5EB391E85CB6A90 CRC64;

Query Match 31.2%; Score 614.5; DB 10; Length 371;
 Best Local Similarity 39.5%; Pred. No. 1.1e-44;
 Matches 139; Conservative 64; Mismatches 126; Indels 23; Gaps 8;
 QY 2 ELQEVLMHNEGEGSTYAKNASTNLAKV KPLFEOCIRELLRANLPINCKIRVADLG 60
 DB 5 DMREFEYMGDGDSTYAKNSLSLOKASDPAKHTLETLELOLYKETRP--KSLGIADLG 61
 QY 61 CASGPNLTLLTVRDIVOSIDKVGQEEKNELERT--IQIFLNDLFQDNFNSVFLLPSFYRL 120
 DB 62 CSSGPNLTLLTVRDIVOSIDKVGQEEKNELERT--IQIFLNDLFQDNFNSVFLLPSFYRL 121
 QY 121 EKENGKTKS-----LISAMPSFGRLFPESNMFLHSCYSVHWLSQVPSGLVIELGIGA 177
 DB 122 KRDNNN--GPCPSVFIAVPPSGFYGRLEFPNTIHFYASHSHMWSKVPYALYDEGKSI 179
 QY 178 NKGSIYSSKGCRRPPVOKAYLDOFTKDTTFLRIHSEKELFSRGMMLTICICK--VDERDEP 235
 DB 180 NKGCVSICSLSSBAVSKAYCSOFKEDFSIFLRCSKENVASAGRMVLIILGREGPDHVRG 239
 QY 236 NPL--DLIDMAINDLIVEGLEEEKLDSFNIPFPSPAEVAVCIYEEBSCETILYETFK 293
 DB 240 NSFWEELLSRSIADIVAOGETEERKLDSDYMHFYAPSADIEGVDKSGSELEERLEME 299
 QY 294 AYDAFSEID-DYVRSHEQIKAEYVASLRSVPEPLIASHFGGAIIMPDLF 344
 DB 300 VKRDKNGTEGDISY-----GKAVAKTVRAVQESMLVQHFGKILDKLF 342

Search completed: July 27, 2002, 05:42:16
 Job time: 492 sec

